Elucidating proton coupling in a vesicular glutamate transporter homolog

Inaugural-Dissertation

zur Erlangung des Doktorgrades

der Mathematisch-Naturwissenschaftlichen Fakultät

der Heinrich-Heine-Universität Düsseldorf

vorgelegt von

Nataliia Dmitrieva

Jülich, 2024

Contents

Publications

In accordance with §6(3) of the doctoral regulations, I state that extracts of this thesis overlap with the following paper:

[A] Dmitrieva N., Gholami S., Alleva C., Alfonso-Prieto M., Carloni P., Fahlke C. Transport mechanism of DgoT, a bacterial homolog of SLC17 organic anion transporters

EMBO Journal; doi: https://doi.org/10.1038/s44318-024-00279-y

Quotations from this publication are indicated throughout the text.

Author's contribution statement: I designed, performed and analyzed all experiments, produced and analyzed classical molecular dynamics simulations data, prepared the figures and drafted the manuscript. The research was done under supervision of Prof. Christoph Fahlke.

List of figures

List of tables

Abstract

Solute carrier 17 (SLC17) family is a group of structurally related proteins that transport organic anions across biological membranes. Members of the family are expressed in the plasma membrane (SLC17A1-4), lysosomes (sialin, or SLC17A5; VNUT, or SLC17A9) and synaptic vesicles (VGLUT1-3, or SLC17A6-8). They exhibit great diversity in substrate specificity and transport mechanism. Together with the other 15 SLC families, they belong to major facilitator superfamily (MFS), the largest superfamily of secondary active transporters. MFS proteins are characterized by their distinct fold with 12 transmembrane helices (TM) that are grouped into two pseudosymmetrical domains. A key element of the transport mechanism is alternating access that involves movement of the two domains around central binding site. Such conformational change interchangeably exposes the binding site to either membrane side and allows substrate transport across the membrane.

To better understand how structurally-similar proteins mediate transport of various substrates with different stoichiometries, we investigated transport mechanism of bacterial homologue, D-galactonate transporter DgoT. Despite moderate sequence identity, DgoT is structurally similar to mammalian SLC17 proteins, particularly VGLUTs and sialin. Two crystal structures in different conformations make DgoT an attractive model system to investigate transport mechanisms in coupled transporters.

We used a combination of computational and experimental approaches to study DgoT. With solid supported membrane-based electrophysiology, we determined that each galactonate molecule is transported together with two protons. Therefore, pH conditions on both sides of the membrane are important; transport can be inhibited by alkaline pHs outside the liposomes due to attenuated proton binding, or by acidic pH inside the vesicles due to impaired proton release. With site-directed mutagenesis, we characterized roles of individual residues in transport activity. Mutations D46N, E133Q, R47Q, R126Q and C43A lead to loss of transport activity, while selective binding of galactonate remains intact. In all mutants, with the exception of R47Q, we observed an electrogenic reaction that can be attributed to a fast substrate binding followed by a conformational change in the protein.

R47 couples substrate binding with the conformational changes in the protein, therefore in the mutant variant only the first step is detected. All-atom molecular dynamics simulations revealed that protonation of D46 and E133 in *apo* DgoT is coupled with rearrangement of gating helices TM1 and TM7 that regulate access to the substrate-binding site from the extracellular side. After gate opening, galactonate can reach the binding site inducing extracellular gate closure and transition to the inward-open conformation, where the substrate is released. Interaction between galactonate and R47 play an important role in this mechanism, and R47Q DgoT fails to close the gate in the presence of the substrate. Using Markov state modeling, we quantified energetics of major conformational changes and demonstrated that the presence of galactonate in the binding site promotes an inward-occluded conformation of the protein, while *apo* transporter favors an outward-open state.

Based on our experimental and computational results, we propose a transport model for galactonate/H$^+$ co-transport in DgoT, which can provide atomic-level insights into the mechanisms underlying coupling in SLC17 proteins.

Zusammenfassung

Die Familie der Solute Carrier 17 (SLC17) ist eine Gruppe von Proteinen, die organische Anionen durch die Membran in Cytoplasma (SLC17A1-4), Lysosomen (SLC17A5 und SLC17A9) oder synaptische Vesikel (SLC17A6-8) transportieren. Sie weisen eine große Vielfalt an Substratspezifität und Transportmechanismen auf. Zusammen mit den anderen 15 SLC-Familien gehören sie zur Major Facilitator Superfamily (MFS). MFS-Proteine zeichnen sich durch ihre ausgeprägte Faltung mit 12 Transmembranhelices (TM) aus, die in zwei pseudosymmetrische Domänen gruppiert sind. Alternating access mechanism umfasst die Bewegung der beiden Domänen um die zentrale Bindestelle, die sie abwechselnd auf beiden Membranseiten freilegt und den Substrattransport ermöglicht.

Hier haben wir den Transportmechanismus eines bakteriellen Homologs, des D-Galactonat-Transporters DgoT, mit einer Kombination aus computergestützten und experimentellen Verfahren untersucht. Durch solid-supported membrane electrophysiology konnten wir feststellen, dass jedes Galaktonatmolekül zusammen mit zwei Protonen transportiert wird. Der aktive Transport kann durch alkalische pH-Werte außerhalb der Liposomen aufgrund einer abgeschwächten Protonenbindung oder durch saure pH-Werte innerhalb der Vesikel aufgrund einer beeinträchtigten Protonenfreisetzung gehemmt werden. Die Mutationen D46N, E133Q, R47Q, R126Q und C43A führen zu einem Verlust der Transportaktivität, während die selektive Bindung von Galaktonat weiterhin beobachtet wird. Bei allen Mutanten, mit Ausnahme von R47Q, beobachteten wir eine elektrogene Reaktion, die auf eine schnelle Substratbindung, gefolgt von einer Konformationsänderung des Proteins, zurückzuführen ist. R47 koppelt die Substratbindung mit der Konformationsänderung des Proteins, daher wird in der mutierten Variante nur der erste Schritt nachgewiesen.

All-atom molecular dynamics simulations ergaben, dass die Protonierung von D46 und E133 in *apo* DgoT mit einer Umstrukturierung des Tores (TM1 und TM7) verbunden ist, die den Zugang zur Substratbindestelle von der extrazellulären Seite aus regulieren. Nach der Toröffnung kann Galaktonat die Bindestelle erreichen, wodurch das extrazelluläre Tor geschlossen wird und in die nach innen geöffnete Konformation übergeht, in der das

Substrat freigesetzt wird. Mithilfe von Markov state modeling haben wir die Energetik der wichtigsten Konformationsänderungen quantifiziert und gezeigt, dass die Anwesenheit von Galaktonat in der Bindestelle eine nach innen geschlossene Konformation des Proteins fördert, während der *apo* Transporter einen nach außen offenen Zustand bevorzugt. Auf der Grundlage unserer experimentellen und computergestützten Ergebnisse schlagen wir ein Transportmodell für den Galaktonat/H$^+$-Kotransport in DgoT vor, das Einblicke auf atomarer Ebene in die der Kopplung in SLC17-Proteinen zugrunde liegenden Mechanismen geben kann.

1 Introduction

1.1 Major facilitator superfamily

Major facilitator superfamily (MFS) is the largest and the most diverse family of secondary active transporters found in all kingdoms of life [1,2]. Members of the family are known to facilitate transport of various substrates across the cell membranes. Structurally, MFS transporters consist of twelve TM helices that form two pseudosymmetrical domains (N- and C-domains, consisting of TM1-6 and TM7-12), which move around the central binding site, opening it to either side of the membrane (Figure 1.1A).

The most extensively studied member of the family is the lactose permease LacY [3–5], the galactoside/H^+ symporter from *Escherichia coli*. It is the first MFS transporter that was purified in a functional state and the first transporter whose high-resolution (3.5 Å) crystal structure was determined [6], providing substantial advances in our understanding of the molecular mechanism of coupled transport. LacY can recognize disaccharides with a D-galactopyranosyl ring and D-galactose, but not D-glucopyranosides or D-glucose [7]. Its canonical substrate is lactose, which binds to the transporter with low affinity (K_d of $\sim 1\,\text{mM}$) [8]. Other high-affinity substrates such as β-d-galactopyranosyl-1-thio-β-d-galactopyranoside (TDG) or 4-nitrophenyl-α-d-galactopyranoside (NPG) were widely used to stabilize different conformations of the protein, which allowed investigation of substrate recognition mechanism and conformational changes in the transporter [6,9–12].

Coupled transport relies on chemical gradients of either substrate or ion. Active transport moves a substrate against its gradient by using energy released from an ionic (or another substrate) gradient (Figure 1.1B). Thus, active transport allows accumulation of a substrate inside the cell driven by the H^+ gradient or vice versa, building an H^+ gradient by moving the substrate downhill its gradient. Alternatively, protonated LacY can perform equilibrium exchange. In this mode, the binding site is exposed to either side of the membrane interchangeably, allowing the substrate to reversibly bind (Figure 1.1C). Unlike active transport, equilibrium exchange does not result in substrate accumulation, since major conformational changes only occur when the substrate is bound, i.e., for each

substrate molecule that is moved across the membrane another substrate molecule must be moved in the opposite direction.
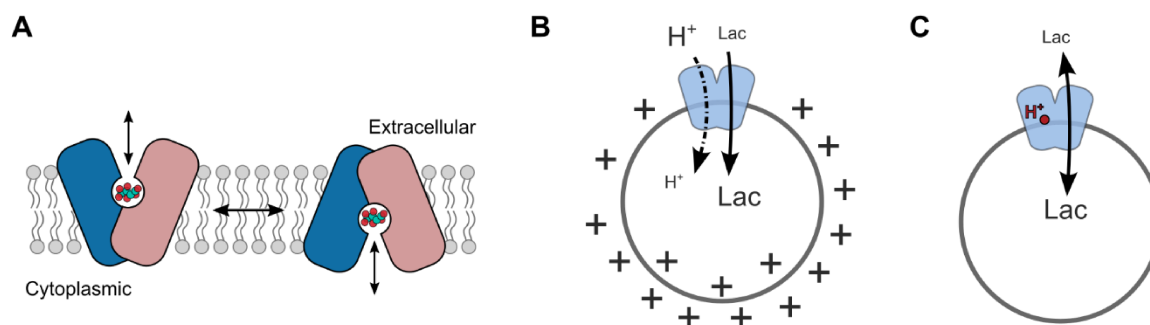
Figure 1.1. Transport reactions mediated by LacY. (A) Schematic representation of the MFS transporter architecture and transport mechanism. (B) Active transport by LacY moves lactose against its gradient using energy stored in H^+ gradient. (C) Equilibrium exchange is facilitated by a protonated transporter, i.e. substrate molecule from one membrane side can be exchanged for another one from the opposite side.

It has been established that of the 417 residues in LacY only few are irreplaceable for active sugar transport [13]: the residues E126, R144, E269 and H322 form the substrate-binding site and R302 and E325 are involved in H^+ transfer. E325 has a pK_a of ~10.5, and at physiological conditions, galactoside binds to the protonated transporter [14]. Notably, mutating the protonation site abolishes coupled transport, without affecting substrate affinity [15]. E325A LacY can isomerize between inward- and outward-facing conformations only when substrate is bound, making the mutant capable of equilibrium exchange, but not active transport [16,17]. R302 is not directly involved in proton transfer, but plays an important role in deprotonation of E325 by lowering its pK_a in inward-facing conformation of the protein [18,19].

1.2 Alternating access mechanism

Unlike channels, transporters do not open simultaneously to both sides of the membrane but switch between different conformations instead. The simple "rocker-switch" model pictures two domains as approximate rigid bodies that are rotating around central cavity, exposing it to either side of the membrane (Figure 1.1A). The first crystal

2

structures of MFS transporters captured proteins in either outward open or inward open state and thus provided evidence for this model. Extensive biochemical experiments on LacY demonstrated that the transporter undergoes major conformational changes [2,5,9,11,20].
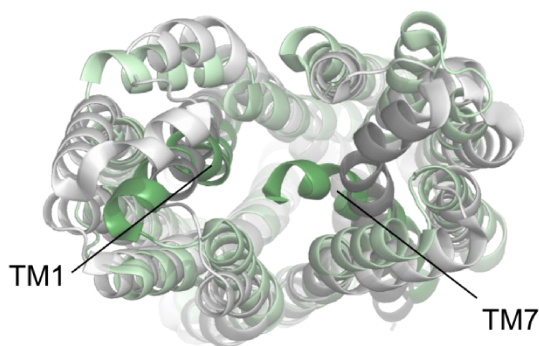
Figure 1.2. Comparison of outward open and occluded states. Aligned structures of outward open conformation of FucP (PDB ID: 3e7q) and outward occluded DgoT (PDB ID: 6e9o) are shown in grey and green, respectively, with gating helices TM1 and TM7 highlighted with darker shade. View from extracellular side.

Later the understanding of the transport mechanisms was refined to account for occluded conformation – an important intermediate state in which binding site is inaccessible from either membrane side [2,21]. In particular, the formation of the occluded conformation involves the bending of gating helices in the presence of a substrate. From the extracellular side, occlusion is mostly achieved by bending of the extracellular ends of TM1 and TM7 (Figure 1.2), – from the opposite side, a similar function is fulfilled by the intracellular ends of TM4 and TM10. This transport mechanism can be described by the more accurate "clamp-and-switch" model, which implies the existence of at least five distinct conformations: outward open, outward-facing occluded, occluded, inward-facing occluded and inward open [1,2]. The switching-step refers to a transition between inward- and outward facing states, which constitutes a rocking motion of the N- and C-domains relative to each other (Figure 1.1A). Whereas the clamping step is characterized as the occlusion of the binding site by gating helices (TM1 and TM7 from extracellular side and TM4 and TM10 from intracellular side) without major changes in protein conformation, i.e., without change to its opposite-facing conformation (Figure 1.2).

Available structural information about different MFS transporters gives a good idea of how key intermediate conformations look. Figure 1.3 shows a comparison of representative structures of proton-coupled transporters captured in different conformations. *Apo* transporters are more likely to be found in open conformations with their binding site exposed to one membrane side. In the presence of a substrate, proteins adopt inward- or outward-occluded conformations, in which the binding site is separated from the bulk by gating helices. The fully occluded conformation is arguably the most challenging structure to capture, since it represents a transition state which is energetically unfavorable [10,22,23]. Conformationally restricted mutants or homologues from different species are widely used to study conformations that are not easily accessible.
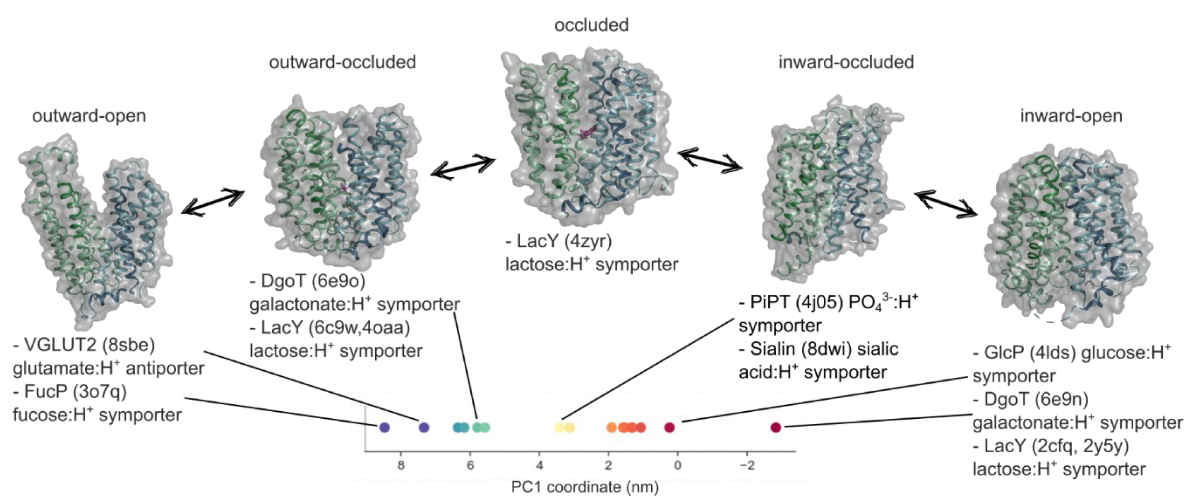
Figure 1.3. Comparison of representative structures of MFS transporters in different conformations. The principal component (PC) analysis from the conserved transmembrane part captures major conformational changes, as represented in projections on the first principal component at the bottom (details of analysis can be found in 2.1.1).

1.3 The solute carrier 17 (SLC17) family

"Solute carrier 17 (SLC17) transporters fulfill a variety of cellular functions [24–26]. They transport diverse anionic substrates; mostly with electrochemical proton (H^+) gradients as the driving force, with a large variability in transport stoichiometry ranging from electrogenic H^+-glutamate exchange by vesicular glutamate transporter (VGLUT) [27]

4

to electroneutral H^+-sialic acid symport by sialin [28,29]." [A] VGLUTs play an important role in signal transmission in the central nervous system [24,30]. Three isoforms (VGLUT1-3) are known in mammals [31,32]. Altered expression is associated with different neurologic diseases such as Alzheimer and Parkinson diseases [33,34]. VGLUTs are complex proteins that facilitate several functions: first described as Na^+-dependent phosphate transporters [35], they were later characterized as chloride-activated glutamate transporters [36–39]. Furthermore, they can act as proton-dependent chloride-channels [38,40]. Sialin is responsible for transport of sialic acids and several additional acidic sugars in lysosomes. Mutations in the transporter cause sialic acid storage disorders such as Salla disease and infantile sialic acid–storage disease [41]. The presence of two closely located glutamate residues in transmembrane helix 4 is required for coupled transport as they participate in proton transfer [28]. Residues surrounding the binding site are responsible for substrate recognition, and their mutation results in inhibition or complete loss of the transport activity [42].

Although a lot of functional and structural information about the SLC17 family has been accumulated over the years, some details about transport mechanisms remain elusive. Despite high structural similarity, substrate specificity and coupling stoichiometry largely differs in the family. Common mechanisms of coupled transport and the ability of similar proteins to fine-tune their function are of great interest.

1.4 Structures of SLC17 family and DgoT

Bacterial D-galactonate/H^+ transporter (DgoT) shares ~20% sequence identity with mammalian members of SLC17 family. Its crystal structures in two different conformations (outward-occluded and inward-open) were determined recently [43], followed by cryo-EM structures of rat VGLUT2 in the outward-open state and human sialin in the inward-facing partially open state [28,44]. Structures reveal common MFS fold with 12 transmembrane helices and are consistent with alternating access mechanism (Figure 1.4).
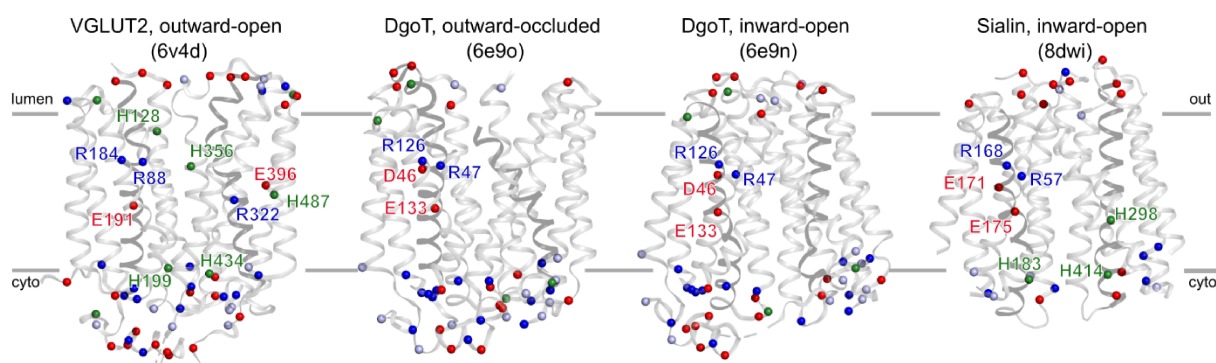
Figure 1.4. Comparison of structures of VGLUT2, DgoT and sialin. Charged residues are shown as spheres and colored according to residue type (arginines/lysins – dark/light blue, histidines – green, aspartates/glutamates – red). The figure was adapted from [A].

Although sequence similarity between DgoT and mammalian SLC17 members is fairly low, structurally they appear to be closely related (RMSD between transmembrane parts of outward-facing DgoT and VGLUT2 structures 4.4 Å). Several residues in the N-terminal domain of DgoT, such as R47, N49, R126 and E133, are conserved throughout the SLC17 family (Figure 1.4, Figure 1.5). While VGLUTs and sialin exhibit titratable residues in the transmembrane region of both N- and C-domains, in DgoT all charged residues are grouped in N-domain, indicating that proton coupling most likely takes place through this part of the protein. For these reasons, DgoT is an attractive model protein to study mechanisms of coupled transport by SLC17 transporters.

Figure 1.5. Partial sequence alignment of DgoT with SLC17 family proteins. Sequences of proteins from *Homo sapiens* (Hs) and *Rattus norvegicus* (Rn) were selected. The numbers indicate the position of residues from DgoT (above) and rat VGLUT2 (below).

DgoT mediates electrogenic transport, indicating that galactonate is transported together with at least two H+, however, the exact stoichiometry has not been determined [43]. D46 and E133 were demonstrated to be essential for transport activity of the protein and therefore were suggested to be proton acceptors [43]. Crystal structures reveal

coordination of galactonate in the binding site by conserved R47 and several polar residues (Y44, Y79, Q164, Q264, S370, N393). Mutation of either of them eliminates transport activity, with the exception of Y44F, which preserves about 25% of WT uptake [42]. These residues were suggested to be important for substrate recognition. Notably, DgoT is highly sensitive with respect to substrate – uptake of ^{14}C-gal by *E. coli* expressing DgoT shows inhibition by 1 mM galactonate but not by 10 mM of its epimer gluconate [43], which differs only in a position of one oxygen atom (Figure 1.6).
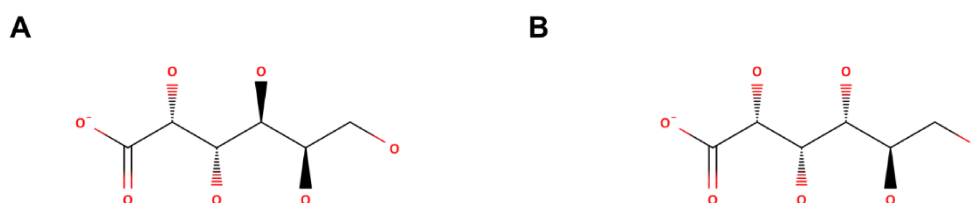
Figure 1.6. Chemical structure of D-galactonate (A) and its epimer gluconate (B).

1.5 Conformational selection and induced fit

Substrate binding is an important part of the MFS proteins' transport cycle. Enzyme catalysis models appeared to be helpful in understanding fundamental mechanisms of protein-substrate interactions [45]. Although transporters do not chemically alter their substrate like enzymes do, the nature of conformational changes accompanying substrate binding is of great interest. MFS transporters bind their substrates via induced fit mechanism, thus employing the energy stored in chemical gradient [46]. Such a model
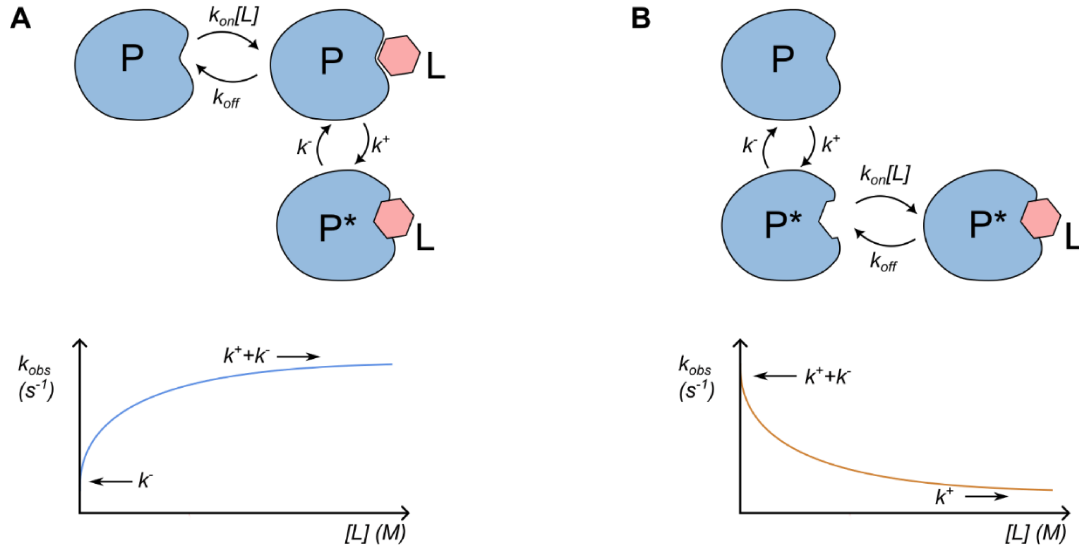
Figure 1.7. Scheme of three-steps binding reactions. (A) Induced fit. (B) Conformational selection. P and P* denote protein in different conformations, L denotes ligand. Expected shape of k_{obs} as function of the ligand concentration for each model is given. Figure was adapted from [47].

postulates that substrate binding promotes formation of occluded state in which affinity is higher (Figure 1.7A).

A model of induced fit mechanism is associated with the following set of differential equations [47]:

$$\begin{pmatrix} d[P]/dt \\ d[PL]/dt \\ d[P^*]/dt \end{pmatrix} = \begin{pmatrix} -k_{on}[L] & k_{off} & 0 \\ k_{on}[L] & -k^+ - k_{off} & k^- \\ 0 & k^+ & -k^- \end{pmatrix} \begin{pmatrix} [P] \\ [PL] \\ [P^*L] \end{pmatrix} \tag{1.1}$$

If conformational changes are much slower than binding and dissociation reactions ($k_{off} + k_{on} \gg k^- + k^+$), the rapid equilibrium approximation is suitable. In this case, the observed rate of the reaction can be expressed as:

$$k_{obs} = k^- + k^+ \frac{[L]}{[L] + K_D} \tag{1.2}$$

Thus, for induced fit mechanism, the experimentally determined observed rate always increases with increasing ligand concentration (Figure 1.7A).

Induced fit is opposed to the conformational selection mechanism, in which protein can isomerize between different states, but the ligand can bind only to specific conformation of the protein (Figure 1.7B). This model can be described with similar set of equations:

8

$$\begin{pmatrix} d[P]/dt \\ d[P^*]/dt \\ d[P^*L]/dt \end{pmatrix} = \begin{pmatrix} -k^+ & k^- & 0 \\ k^+ & -k^- - k_{on}[L] & k_{off} \\ 0 & k_{on}[L] & k_{off} \end{pmatrix} \begin{pmatrix} [P] \\ [P^*] \\ [P^*L] \end{pmatrix} \tag{1.3}$$

In rapid equilibrium approximation, observed rate of the reaction can be expressed as:

$$k_{obs} = k^+ + k^- \frac{K_D}{[L] + K_D} \tag{1.4}$$

In both cases, dependency of observed rate on ligand concentration allows for the determination of 3 parameters. However, it can be challenging to experimentally determine which mechanism is involved in the measured process. While in the rapid equilibrium approximation, conformational selection can be easily recognized by a decrease in k_{obs} with increased substrate concentration, in a scenario with $k^- + k^+ > k_{off} > k^+$, eq. (1.4) cannot be applied [47]. Instead, k_{obs} would increase with substrate concentration, much like in case of induced fit mechanism.

Moreover, in a real system, a mixed mechanism is possible, which contains both induced fit and conformational selection pathways [48]. Thus, an inference as to which mechanism is observed in the experiment, must be done with cation, especially if the observed rates increase with increasing ligand concentration.

1.6 Aims of the thesis

This works aims to decipher molecular determinants of coupled transport in SLC17 family through atomic-level understanding of galactonate transport by a bacterial homologue DgoT. Using the combination of computational approaches, we address the following aspects of the transport mechanism:

- Determination of galactonate:H$^+$ symport stoichiometry
- Effect of pH on transport activity
- Role of protonation in extracellular gate dynamics
- Coupling between substrate binding and conformational changes
- Determinants of substrate release
- Energetics of major conformational changes

- Identification of residues that are crucial for transport and characterization of their roles

Based on obtained information about the main intermediate steps of the transport cycle, we propose a detailed scheme of galactonate:H^+ symport by DgoT.

2 Methods

2.1 Computational methods

Recent progress in structural biology, as well as computational methods of protein structure prediction, led to emergence of increasing number of reliable three-dimensional models of biomolecules. However, it remains particularly challenging to obtain structures of transient conformations since they are usually energetically unfavorable. A possible solution is comparison of homologous protein structures that are captured in different conformations. This allows to identify key structural elements and build a common transport cycle model.

Although structural information provides important insights into molecular mechanisms of protein function, static structures cannot capture the protein dynamics. To address this issue, molecular dynamics (MD) simulations are commonly used as an instrument to investigate biomolecular processes at atomic level [49].

2.1.1 Principal component analysis

Principal component analysis (PCA) was used to compare crystal structures of DgoT with available structures of related MFS proteins [50,51]. Two 200-ns long trajectories with *apo* DgoT in inward and outward conformations (started with respective crystal structure) were artificially stitched together to analyze the motions that contribute most to major conformational change. A covariance matrix was built with the coordinates of backbone atoms of the following residues from transmembrane part of DgoT: 28-55, 65-92, 95-114, 122-146, 152-181, 183-206, 256-276, 294-318, 325-343, 351-366, 369-380, 390-414, 416-437. After diagonalizing the covariance matrix, the eigenvectors and corresponding eigenvalues were extracted. The first eigenvector that represents the biggest correlated linear motion in the system captured about 75% of all motions in the selected atom group. Therefore, the first eigenvector was chosen as an appropriate coordinate that can describe protein conformation as inward-facing, outward-facing or occluded.

11

All structures of interest were projected onto the first eigenvector. In addition to DgoT, analysis included a few functionally (proton-coupled symporters) or structurally (VGLUT and sialin) similar transporters (Table 2.1).

PDB ID	Name	Function	Conformation	Projection on first eigenvector (nm)
6e9n	DgoT	H^+/galactonate symport	Inward open	-2.811
6e9o			Outward occluded	7.514
3o7q	FucP	H^+/fucose symport	Outward open	8.475
4j05	PiPT	H^+/PO_4^{3-} symport	Inward occluded	3.419
4lds	GlcP	H^+/glucose symport	Inward open	0.248
4oaa	LacY	H^+/lactose symport	Outward occluded	6.351
6c9w			Outward occluded	5.575
4zyr			Occluded	6.175
2cfq			Inward open	-2.886
6v4d	VGLUT2	H^+/glutamate antiport	Outward open	7.336
8dwi	Sialin	H^+/sialic acid symport	Inward occluded	-1.708

Table 2.1. List of structures used for principal component analysis.

Larger values of projection on the first eigenvector corresponded to outward conformations, while smaller values were obtained for inward conformations. Thus, PCA was able to correctly identify conformational state of a transporter and confirmed that two crystal structures of DgoT correspond to two significantly different conformations.

2.1.2 Unbiased MD simulations

"DgoT protein coordinates obtained from the Protein Data Bank (inward-facing, PDB ID: 6E9N; outward-facing galactonate-bound, PDB ID: 6E9O [43]) were used as the starting coordinates for MD simulations. The D46 and E133 protonation state and substrate occupation were modified as described in the text. Standard protonation states at neutral

pH were assigned to all other residues (deprotonated aspartate and glutamate residues, singly protonated histidine residues) except for H56, which forms a salt bridge with E180 and, therefore, was set as doubly protonated. Missing residues (235–242 in the inward-facing structure, 231–243 and 277–290 in the outward-facing structure) were modeled using the SWISS MODEL server [52] with a final residue range of 27–442 used for both protein structures. The N- and C-termini were capped with neutral acetyl and methylamide groups, respectively. For modeling DgoT single point mutations, we used PyMOL [53].

The initial protein orientation within the membrane was set to the corresponding DgoT structure available in the Orientations of Proteins in Membranes database [54]. The protein was then embedded into a phosphatidylcholine (POPC) bilayer using *g_membed* [55] in GROMACS and solvated in a box with dimensions ~120 × 120 × 100 Å, which was chosen to ensure a minimum distance between periodic copies of at least 30 Å. The protein/membrane system was surrounded by a ~100 mM solution of Na^+ and Cl^- ions. Ions were described using default CHARMM parameters, and the CHARMM TIP3P model was used for water molecules. MD simulations were performed using the GROMACS software package (versions 2018, 2020, and 2022) [56] with a CHARMM36m force field[57]. Galactonate parameters were obtained using the SwissParam server [58] and added to the forcefield. An integration time step of 2 fs was used. Van der Waals interactions were calculated with the Lennard–Jones potential and a cutoff radius of 1.2 nm, with forces smoothly switched to zero in the range of 1.0–1.2 nm and no dispersion correction applied. Electrostatic interactions were calculated by the particle mesh Ewald method [59], with a real-space cutoff distance of 1.2 nm. All simulations were done in an isothermal–isobaric ensemble, with the temperature set to 310 K using a v-rescale thermostat [60] and a time constant of 0.5 ps. The thermostat was applied separately to the protein, lipid bilayer, and aqueous solution containing ions. The same groups were used for the removal of the center-of-mass linear motion.

The protein was equilibrated in three steps using the velocity-rescale thermostat and Berendsen [61] pressure coupling. The first step lasted 50 ns and was run with positional restraints on protein atoms with a harmonic potential with a force constant of 1000 kJ mol^{-1} nm^{-2} to allow for equilibration of water and ions. In the second step, only the

backbone atoms of the protein were restrained to enable side chains to equilibrate. Lastly, the system was equilibrated for 1 ns without positional restraints to obtain the velocities used in the following production runs. Production MD simulations used Parrinello–Rahman [62] pressure coupling in a semi-isotropic manner with a time constant of 0.5 ps." [A] Summary of all simulated systems and length of trajectories can be found in Table 7.1.
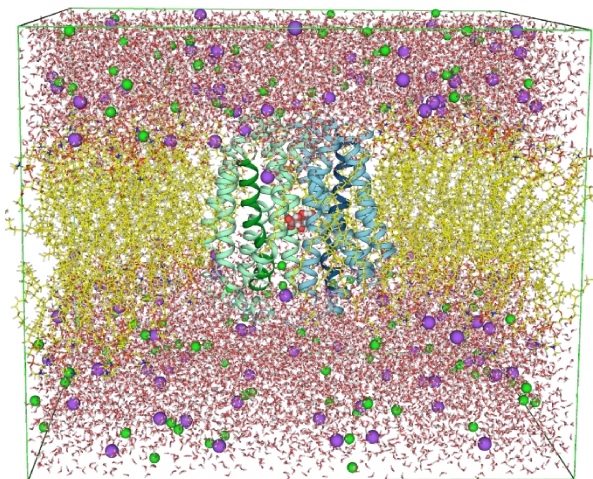
Figure 2.1. Molecular dynamics simulation box. DgoT monomer (blue and green helices) was embedded in POPC bilayer (yellow sticks) and soaked in water (red-white sticks) and NaCl (green and purple spheres). Bound galactonate molecule in the binding site is shown as grey and red spheres.

2.1.3 Time-lagged independent component analysis

While unbiased MD simulations are a powerful tool for exploration of protein dynamics, slow processes such as conformational changes often cannot be observed in a timescale of a single trajectory. The complexity of the system presents an additional challenge, as the choice of appropriate reaction coordinate that describes accurately protein conformation is not always trivial.

Availability of crystal structures of DgoT in two distinct conformations simplified the exploration of conformational space. To describe major conformational changes, we focused on the movement of N- and C-terminal domains relative to each other. We selected every 5th Cα atom of the transmembrane part of the protein and computed the distances between each selected atom located on the N-terminal (TM1–6) and C-terminal (TM7–12) domains, respectively (Figure 2.2A). This resulted in 27 × 26 = 2809 pairwise distances for

each frame. To reduce dimensionality, time-lagged independent component analysis (TICA) [63] with a lag time of 50 ns was applied on a data collected from initial round of unbiased simulations. The first two components (ICs) obtained were kept as they had slower timescales than the other independent components (Figure 3.9B, C). The first IC is mainly responsible for separation of inward- and outward-facing states, as conformations captured in two crystal structures are well divided along it. The second IC captures dynamics of outward-facing DgoT and separates outward open and outward occluded conformations (Figure 2.2B).
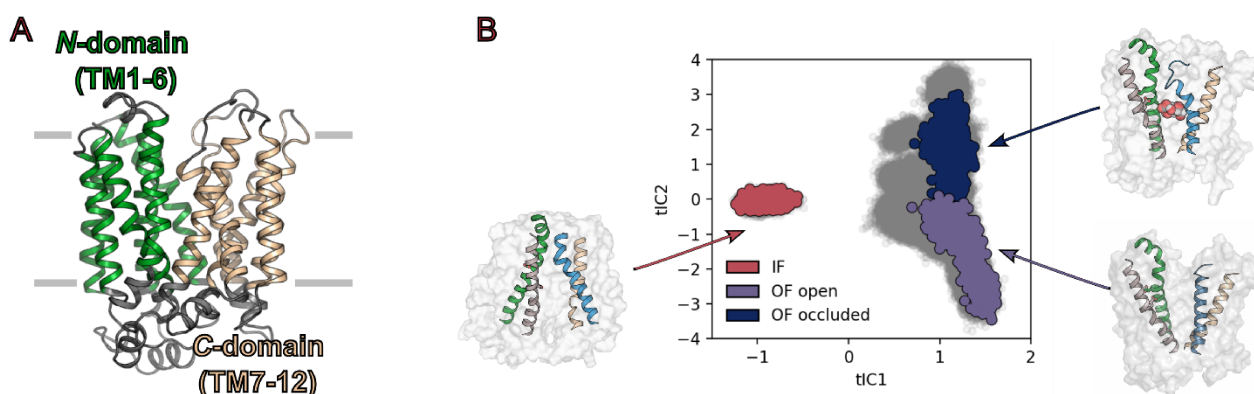
Figure 2.2. TICA. (A) Outward-facing DgoT structure (PDB ID 6E9O) with transmembrane part of N- and C-terminal domains colored as green and wheat, respectively. (B) Conformations sampled in initial round of unbiased simulations projected on first TICA components (grey dots). Data from representative trajectories, in which only inward-facing, outward open or outward occluded conformations were observed, are shown as colored points. Illustrative protein structures are added next to each of three mention states.

To fully sample the transitions between all relevant intermediate states, we focused on two systems: (i) galactonate-bound DgoT with protonated D46 and E133 (responsible for translocation of the substrates across the membrane) and (ii) apo DgoT with deprotonated D46 and E133 (responsible for reorientation of the empty transporter after substrate release). For them additional unbiased simulations 500-800 ns long were performed, making up total simulation time ~25 μs for substrate-bound system and ~37 μs for *apo* system.

15

2.1.4 Markov state models

Markov state models (MSMs) were constructed using PyEMMA 2 software [64] from multiple unbiased MD trajectories 400-800 ns long (Figure 3.9A). The acquired data was projected onto the first two tICA reaction coordinates. The decomposed data was clustered into 200 discrete states using the k-means algorithm. The implied timescales plot demonstrates the Markovian behavior after a lag time of ~ 80 ns (Figure 3.9B, C) for both systems. For constructing the MSM, a lag time of 100 ns was chosen from the implied timescales plot. For model validation, we performed a Chapman–Kolmogorov test (Figure 3.9D, E). Metastable states were identified using Perron-cluster cluster analysis (PCCA)[65]. For each system, the implied timescale plots (Figure 3.9B, C) revealed two slow processes; therefore, we clustered the microstates into three metastable states.

2.1.5 Calculation of electrostatic potential

A GROMACS-based tool *g_elpot* [66] was used for quantification of electrostatics in unbiased simulations; source code, installation instructions, and usage recommendations can be found at https://jugit.fz-juelich.de/computational-neurophysiology/g_elpot. "The distribution of electrostatic potential was calculated via the smooth particle mesh Ewald (SPME) method. For our system, the SPME potential was calculated on a grid of 256 × 256 × 208 points with an inverse Gaussian width β of 20 nm-1." [A] *g_elpot* calculates distribution of potential within water molecules present in the system. Time-resolved electrostatic potentials can be calculated in arbitrary regions of the system. In practice, average potential of water molecules within user-defined radius of chosen point is calculated, along with the total number of water molecules. Dehydrated frames can later be discarded from the analysis, leaving only information about the potential in the region of interest.

Alternatively, *g_elpot* can be used to calculate potential within an arbitrary atom group, e.g., at the tip of residue [67]. In this case, some insights about electrostatics in the system can be gain independently on the hydration of the region. A group constituting of a few atoms with fixed geometry is used, and electrostatic potential if an average of the SPME

potential in a 0.15-nm sphere around the center of the geometry of the group. Unlike with water-based electrostatic potential, in this application the absolute potential value can only be compared with that of residues of the same type or with that of the same residue under different simulation conditions.

"We exploited the extended functionality of the tool [67] to calculate the potential for the carboxyl groups of protonated D46 and E133. The electrostatic potential is an average of the SPME potential in a 0.15-nm sphere around the following atom groups (CHARMM naming convention): OD1, OD2, and CG of D46 and OE1, OE2, and CD of E133. Since two selected groups are chemically identical, the residue-specific short-range part of the potential is the same and, therefore, comparison of the potential values is justified. The time course of the electrostatic potential was averaged across each trajectory." [A]

2.2 Experimental methods

2.2.1 D-galactonate preparation

Crystalline Na$^+$ D-galactonate was prepared from commercially available calcium salt according to previously reported method [43,68,69]. 10 g Ca^{2+} D-galactonate (Biosynth) was resuspended in 100 ml water. An equimolar amount of oxalic acid dihydrate (2.93 g) was added to the mixture and stirred for 10 minutes at 55 °C. The precipitated calcium oxalate was then separated from aqueous D-galactonic acid by filtration under vacuum through 0.22 μm nylon. Sodium hydroxide was titrated into the solution to pH 7. Absolute ethanol was added in a 3:1 (v/v) ratio, the mixture stored at 4 °C overnight, and the resulting crystalline precipitate removed by filtration and washed with absolute ethanol. The precipitate was dried for 24 hours in a vacuum desiccator. The resulting crystalline Na$^+$ D-galactonate was then stored at room temperature for subsequent use.

2.2.2 Protein expression and purification

"The full-length DgoT gene (GenBank accession number AKK15832.2) was subcloned into a pQE60 vector through the NcoI and HindIII restriction sites in fusion with a C-terminal thrombin cleavage site and decahistidine tag. Mutant constructs were generated using PCR-based mutagenesis and verified by DNA sequencing. Protein expression and purification were performed using a published procedure, with modifications [43]. *E. coli* C41 cells transformed with pQE60 DgoT WT were grown at 37 °C in TB medium supplemented with 2 mM $MgSO_4$. When an OD_{600} of 0.6–0.8 was reached, gene expression was induced with 1 mM IPTG and cells were grown for a further 4 h (typical yield of 15 g per 1 L culture). After sedimentation, cells were flash frozen in liquid nitrogen and stored at -80 °C for later use, Next, cells (15 g) were resuspended in 20 mM Tris (pH 7.4) and 300 mM NaCl (50 mL volume) containing complete protease inhibitor cocktail (Roche) and lysed by sonication. Debris was removed by centrifugation at 12,000 × *g* for 15 min, and membranes were collected at 200,000 × *g* for 1 h, flash frozen in liquid nitrogen, and stored at -80 °C until use.

A frozen membrane pellet from 15 g cells was resuspended in 20 mL membrane buffer (20 mM Tris (pH 7.4), 150 mM NaCl) containing cOmplete protease inhibitor cocktail, using a glass Dounce homogenizer, and then *n*-dodecyl-D-maltoside (DDM) was added to 1.4%. Membranes were solubilized for 2 h at 4 °C and the insoluble fraction was removed by ultracentrifugation at 75,000 g for 30 min at 4 °C. The supernatant was diluted 1:2 with solubilization buffer, imidazole was added to 15 mM, and pH was adjusted to 7.8–8.0. CoNTA resin (3 mL) was washed with 10 column volumes (CV) of membrane buffer, added to the supernatant, and incubated at 4 °C for 1.5 h under gentle agitation. The resin was then washed with 10 CV of wash1 buffer (20 mM Tris, 150 mM NaCl, 15 mM imidazole, 0.1% DDM, pH 8.0) and 10 CV of wash2 buffer (20 mM Tris, 150 mM NaCl, 0.1% lauryl maltose neopentyl glycol (LMNG), pH 7.4). Protein was eluted with 4 CV of elution buffer (20 mM Tris, 150 mM NaCl, 0.1% LMNG, 150 mM imidazole, pH 7.4) into 0.5 CV fractions, and 5 mM EDTA was added immediately after collection. Protein concentration in the eluted fractions was estimated by measuring the absorbance at 280 nm (NanoDrop), and fractions with the highest protein concentrations were combined. Imidazole was removed

using a PD-10 desalting column, with protein eluted with buffer (20 mM Tris, 150 mM NaCl, 1 mM EDTA, 0.05% LMNG, pH 7.4) and stored overnight at 4 °C.

The purified protein was concentrated to 5–8 mg/ml using a 50 kDa molecular weight cutoff (MWCO) centrifuge concentrator (Millipore) and loaded in 0.5 mL portions onto a Superdex 200 Increase 10/300 GL column (GE Healthcare Life Sciences) preequilibrated with size-exclusion chromatography (SEC) buffer (20 mM HEPES, 150 mM NaCl, 0.05% LMNG, pH 7.4). The peak fractions were combined, flash frozen in liquid nitrogen, and stored at -80 °C until use." [A]

2.2.3 Reconstitution of proteoliposomes

"*E. coli* polar lipids (Avanti Polar Lipids, *E.coli* polar lipid extract, 25 mg/ml solution in chloroform) were dried under nitrogen and then under vacuum overnight. The dried lipid film was resuspended to a lipid concentration of 10 mg/ml by stirring in reconstitution buffer (1 mM HEPES pH 7.4, 150 mM NaCl, 2 mM MgSO$_4$) for 1 h at room temperature (RT). After the lipids were completely dissolved, the suspension was frozen in liquid nitrogen, stirred for another 1 h at RT, and sonicated using a UP50H ultrasonic processor equipped with a microtip until clear (2 or 3 cycles of 30 s). The formed liposomes were destabilized by adding 0.6% Triton X-100 and incubated for 45 min at RT under gentle agitation. Purified DgoT was added to the destabilized liposomes at a LPR of 5, 10, or 20 and then incubated for 1 h at 4 °C. To remove the detergent, 150 mg SM2 Bio-beads was added per 1 mL liposome suspension. After incubation for 1 h at 4 °C, another 150 mg Bio-beads was added per 1 mL suspension, followed by another 1 h incubation at 4 °C, and then the beads were removed using a disposable column. A third volume of Bio-beads was then added (400 mg per 1 mL suspension), incubated overnight at 4 °C, and removed using a disposable column. Empty liposomes were prepared in parallel using the same procedure but with no added protein." [A]

The efficiency of reconstitution was evaluated by centrifugation of liposomes in sucrose gradient. Clear centrifuge tube was filled with 4 sucrose layers, from bottom: 40%, 30%, 20% and 10% sucrose in reconstitution buffer (w/v), 0.75 mL each. Liposome sample with

19

protein concentration of ~ 1 mg/mL was added on top of the 10% sucrose layer, volume of added sample was 70 μL. Centrifugation was performed for 3 h in swinging-bucket rotor SW 60 Ti (Beckman Coulter) at 50,000 rpm (approximately 250,000 g). After the centrifugation 11 layers were carefully taken from top of the tube without mixing, volumes of samples were equal except for the most bottom layer that had slightly larger volume due to tube geometry. All collected layers were subjected to SDS-PAGE to estimate fraction of protein that was successfully reconstituted.

2.2.4 SDS-PAGE and western blot

Sodium dodecyl sulfate–polyacrylamide gel electrophoresis (SDS-PAGE) was used to analyze protein samples. Gels were prepared as described in Table 2.2, run in vertical electrophoresis cell (BioRad) with SDS-running buffer (25 mM Tris base, 192 mM glycine, 0.1% SDS) at 180V for 40-50 min. Samples were supplemented with Laemmli Sample Buffer (5x stock: 300 mM Tris pH 6.8, 10% SDS, 50% glycerol, 50 mM DTT, 0.05% Bromphenol blue) and added to gel immediately or stored at -20 °C until further use.

10% separating gel		4% stacking gel	
Substance	Volume (total 10 ml)	Substance	Volume (total 10 ml)
H$_2$O	4.0 ml	H$_2$O	6.1 ml
Tris 1.5M pH 8.8	2.5 ml	Tris 1.0M pH 6.8	2.5 ml
Acrylamide	3.3 ml	Acrylamide	1.3 ml
SDS 10%	100 μl	SDS 10%	100 μl
APS 10%	100 μl	APS 10%	140 μl
TEMED	10 μl	TEMED	12 μl

Table 2.2. SDS-PAGE gel composition. Acrylamide stands for Acrylamide/Bis Solution, 37.5:1 (30% w/v)

Gel with separated proteins was either stained or subjected for transfer for western blot (WB). To estimate the size of proteins in sample molecular weight standard PageRuler™ Plus (Thermo Fisher, Cat. 26619) was used as a reference.

Staining: gel was incubated in staining solution (for 1L: 1.5 g Coomassie Brilliant Blue (SERVA), 450 ml ethanol denatured, 92 ml acetic acid (96%)) for 0.5-1 h, followed by incubation in destaining solution (for 1L: 300 ml ethanol denatured, 30 ml acetic acid (96%)) ON.

WB: gel was briefly rinsed in transfer buffer (25 mM Tris base, 192 mM glycine, pH 8.3; buffer was freshly prepared from 10x stock solution), placed in "transfer sandwich" (fiber pad, filter paper, nitrocellulose membrane, gel, filter paper, fiber pad; all components equilibrated in transfer buffer) and fixed in the chamber vertically. Transfer buffer was used to fill up the chamber. Ice pack and magnetic stirrer were placed in the chamber to avoid overheating of the gel during the procedure. The transfer was performed for 50 min under constant voltage of 90 V, current was limited to 350 mA.

Transferred proteins on the membrane were blocked with PBS-T buffer (140 mM NaCl, 2.68 mM KCl, 9.89 mM Na_2HPO4 x 2 H_2O, 1.76 mM KH_2PO_4, pH 7.3, 0.1% Tween20) supplemented with 2% non-fat milk for 1 h at RT, incubated with primary antibodies (Penta-His Mouse Antibody BSA-free, Qiagen) dissolved 1:1000 in PBS-T buffer with 2% milk for 2.5-3.5 h, rinsed 5 times with PBS-T buffer and incubated with secondary antibodies (Anti-Mouse IgG (Fc specific)–Alkaline Phosphatase antibody produced in goat, Merck) dissolved 1:1000 in PBS-T buffer with 2% milk for another 40 min. Then membrane was thoroughly washed with PBS-T buffer to remove any remaining non-bound antibodies. To visualize protein bands membrane was incubated with 10 ml phosphatase buffer (100 mM NaCl, 5 mM $MgCl_2$, 100 mM Tris, pH 9.5) supplemented with 33 µl BCIP (5-Bromo-4-chloro-3-indolyl phosphate) and 66 µl NBT (nitro blue tetrazolium) for several minutes and then thoroughly washed with distilled water.

2.2.5 pH electrode-based assays

pH electrode-based assays were used to monitor coupled transport in bacterial cells overexpressing DgoT and to investigate binding order in purified transporter. Experimental procedures were performed using modified protocols from [12,70]. "DgoT was expressed in *E. coli* C41 cells using the same procedure as for purification. After expression, bacteria

were pelleted, washed twice in assay solution containing 250 mM KCl, 1 mM MgSO$_4$, 2 mM CaCl$_2$ and resuspended in assay solution to reach bacterial density of OD$_{600}$ 15. All centrifugations were performed at 2,200 g, 7-8 min, cells were resuspended on vortex. Substrates (galactonate or gluconate) that were assayed were dissolved in assay solution at a concentration of 160 mM and the pH was adjusted to 6.5-6.7 with KOH, being always lower than pH of bacterial suspension in each experiment. 800 μl of bacteria suspension were transferred into a 2 ml reaction tube (Eppendorf) and the pH value of the bacterial suspension (extracellular medium) was measured using a micro pH electrode with integrated temperature sensor (Xylem, SI Analytics) under continuous stirring. The pH of the bacterial suspension was adjusted to pH 6.7 using KOH and HCl. After recording of baseline for 60 seconds 50 μl of the corresponding compound was added (to reach final concentration of 10 mM) to the bacterial suspension. Galactonate or gluconate were dissolved in assay solution with concentration of 160 mM, and pH was adjusted to be 6.5-6.7 before the experiment, always lower than pH of cells suspension. All experiments were performed at RT (20-21 °C).

Experiments with purified protein were performed similarly. Purified DgoT (stored at -80 °C) was thawed, SEC buffer (used for storage of purified protein) was replaced using PD-10 desalting column with assay solution containing 150 mM NaCl, 0.05% LMNG, pH ~7.4." [A]

2.2.6 SSM-based electrophysiology

Solid-supported membrane-based electrophysiology (SSME) is a technique developed for the studies of transporters that are difficult to investigate using patch-clamp [71]. The principle of the method is schematically illustrated on Figure 2.3. SSM is formed by addition of lipid on gold-coated sensor chip with an alkylmercaptane monolayer on top. Membrane preparation (for example, proteoliposomes with reconstituted protein or membrane vesicles) with protein of interest can be added on SSM. This leads to the stable adsorption of the membranes to the SSM and the formation of a capacitively coupled compound membrane (Figure 2.3B). Measuring electrode is the sensor chip and reference electrode is

inserted in the solution (Figure 2.3A). The sensor is placed inside a Faraday cage and connected to a fluidic system which allows fast solution exchange.
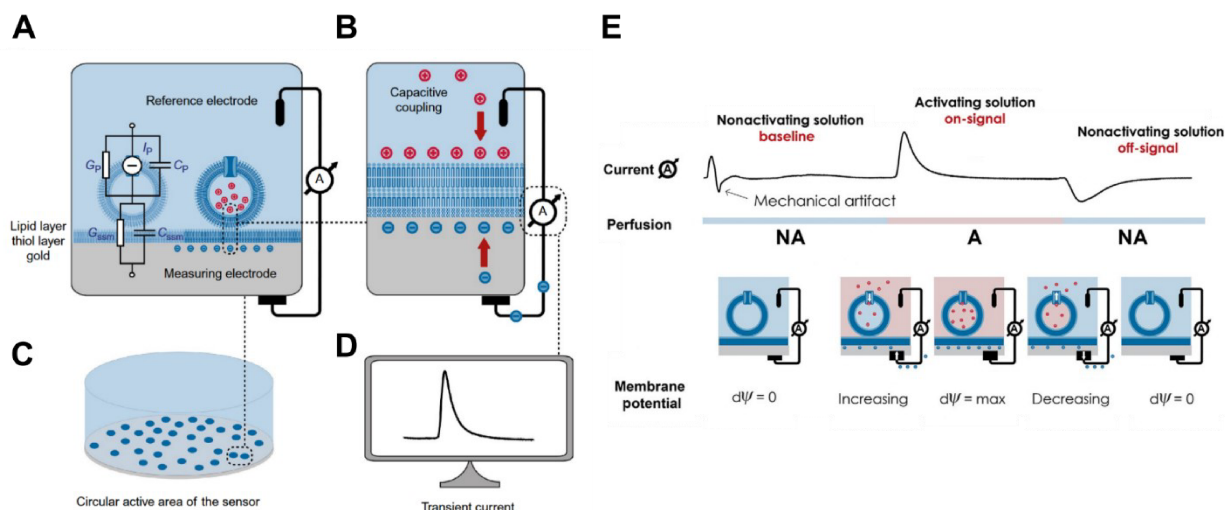
Figure 2.3. SSM-based electrophysiology. (A) Equivalent electrical scheme of SSM with adsorbed proteoliposomes with reconstituted protein of interest. (B) Expansion of the interface of the SSM and the proteoliposomal membrane. (C) Proteoliposomes adsorbed on SSM. (D) Currents measured in SSME are transient due to capacitive coupling. (E) Scheme of single solution exchange protocol. The figure was adapted from [12].

All measurements in SSME are done at 0 mV, instead, the driving force for the electrogenic transport is substrate concentration gradient. Transport of charged substrates or ions into the liposomes or vesicles generates a membrane potential (Figure 2.3E). This potential is detected via capacitive coupling between the membrane and the SSM on the gold layer of the sensor. At some point the membrane potential equals the chemical driving force and the transport process comes to a halt. Therefore, any current measured with SSME is transient (Figure 2.3D). Since the current decay is fast, one measurement takes only one second. Due to the high stability of the SSM, multiple measurements can be performed using the same sensor and different buffer conditions to determine kinetic parameters.

"Gold electrode sensors (1 or 3 mm) were prepared as previously described [71]. Briefly, sensors were incubated for at least 30 min in an octadecane thiol solution, and then rinsed thoroughly with isopropanol and water. The solid-supported membrane (SSM) was prepared by pipetting 1.5 μL diphytanoyl phosphatidylcholine dissolved in n-decane onto

the electrode surface, followed by 100 μL aqueous buffer. Immediately prior to measurements, liposome samples were thawed, diluted to a final lipid concentration of 1 mg/mL and briefly sonicated. Each liposome sample (10 μL) was pipetted onto a SSM sensor and adsorbed by centrifugation at 2,200 × g for 30 min at RT. All experiments were repeated for at least three sensors, with each condition measured at least twice. All solutions were buffered in 100 mM potassium phosphate (KPi) for each pH used.

For measurements with a single solution exchange protocol, three phases of 1 s duration were applied: flow of nonactivating (NA) solution, activating (A) solution, and NA solution. Only the A solution contained galactonate. In experiments with variable pH, pH of NA and A solutions was kept constant within a single experiment. Between experiments using different pH values, the sensor was incubated at the new pH for 5 min to equilibrate the intraliposomal pH. In experiments with WT, D46N, E133Q, and R126Q DgoT, the NA solution contained gluconate to compensate for galactonate in the A solution. For R47Q DgoT, glutamate was used in NA solution instead, allowing a direct comparison of the responses to galactonate and gluconate applications.

To determine the apparent pK values, normalized peak currents (measured using 3 mm sensors) were fitted with one of the following equations:

acidic deactivation (pH ≤ 8):

$$I_{norm}(pH) = \frac{I_{max}}{1 + 10^{pK_a - pH}};\qquad\qquad(2.1)$$

alkaline deactivation (pH ≥ 7.5):

$$I_{norm}(pH) = \frac{I_{max}}{1 + 10^{pH - pK_b}}.\qquad\qquad(2.2)$$

Currents were normalized to the I_{max} value obtained by the fit of peak currents measured on the same sensor.

For the analysis of pre-steady state currents, 1 mm sensors were used. Rate constants for the observed charge displacements ($k_{obs} = 1/\tau$) were derived from the transient currents

by fitting the decay with a monoexponential function $I = A * \exp(-t/\tau)$. This two-step reversible reaction can be described as follows:

$$P + S \leftrightarrow PS \underset{k^-}{\overset{k^+}{\rightleftarrows}} P^*S, \qquad (2.3)$$

where P is the protein, S is the substrate, and P^* is the protein after conformational changes. The first step is the binding reaction described by the dissociation constant K_D, and the second step is the substrate-induced conformational change characterized by the forward and reverse rate constants. Assuming that substrate binding is rapid, the observed rate constant has hyperbolic dependence on the substrate concentration [72]:

$$k_{obs} = k^- + k^+ \frac{[gal]}{[gal] + K_D}. \qquad (2.4)$$

Measurements under asymmetrical pH conditions were done using a double solution exchange configuration. A resting (R) solution phase of 1 s and incubation period of 5–20 min was added to the beginning of each measurement to allow the intraliposomal pH to adjust to the pH of the R solution. Afterwards, a normal single solution exchange protocol was used to establish the pH gradient (during NA phase) and substrate gradient (A phase).

To determine the transport stoichiometry, we used a reversal assay as previously described [73], with a single solution exchange protocol comprising three phases (NA, A, and NA) extended to 2, 3, and 3 s, respectively. Between measurements, the sensor was rinsed 3–5 times with NA solution; the current responses were recorded and used as a baseline. The same protocol was used for samples with empty liposomes to account for solution exchange artifacts. The entire A phase was integrated to obtain the transported charge values. After subtraction of the negative control (integrated current recorded with empty liposomes), the transported charge values were used to determine the transport stoichiometry. The internal (NA) solution used for this experiment had a pH of 7.3 and contained 0.5 mM galactonate and 7.5 mM gluconate, and applied external (A) solutions had a pH of 7.6 and contained X mM galactonate and 8–X mM gluconate. This setup resulted in dµ $_{gal}$/ dµ $_{H+}$ values of 0–4." [A]

Transport stoichiometry was determined using the same rationale as used in [73]. Coupled transport can be described by the substrate and ion concentrations:

$$nIon_{out} + mSubstrate_{out} \rightleftharpoons nIon_{in} + mSubstrate_{in}. \qquad (2.5)$$

In SSME, no voltage is applied to the membrane, therefore chemical potential of ion and substrate across the membrane is defined only by their concentrations:

$$\Delta\mu_i = RT \ln\left(\frac{[Ion]_{in}}{[Ion]_{out}}\right), \qquad (2.6)$$

$$\Delta\mu_s = RT \ln\left(\frac{[Substrate]_{in}}{[Substrate]_{out}}\right). \qquad (2.7)$$

The free energy for the coupled transport reaction is given by:

$$\Delta G = n\Delta\mu_i + m\Delta\mu_s. \qquad (2.8)$$

When current reverses direction, net transported charge is 0 and $\Delta G = 0$. Under these conditions:

$$-\frac{n}{m} = \frac{\Delta\mu_s}{\Delta\mu_i} = \frac{RT \ln\left(\frac{[Substrate]_{in}}{[Substrate]_{out}}\right)}{RT \ln\left(\frac{[Ion]_{in}}{[Ion]_{out}}\right)}, \qquad (2.9)$$

or

$$\frac{n}{m} = \frac{RT \ln\left(\frac{[galactonate]_{in}}{[galactonate]_{out}}\right)}{RT \ln\left(\frac{[H^+]_{out}}{[H^+]_{in}}\right)}. \qquad (2.10)$$

Thus, the plot of transported charges against $\Delta\mu_{gal}/\Delta\mu_{H^+}$ exhibits an x-intercept at n/m, which provides determination of the transport stoichiometry.

2.2.7 NanoDSF

Nano-differential scanning fluorimetry (nanoDSF) is a technique that enables monitoring conformational stability of the sample by measuring changes in intrinsic protein fluorescence [74,75]. Protein in the sample is slowly heated which leads to its unfolding. As a result, emission spectrum peak of tryptophan residues in the protein shifts from 330 nm to 350 nm. Plot of ratio F_{350}/F_{330} reports on fraction of unfolded protein at each temperature, and its inflection point is at the temperature at which half of the protein in sample is unfolded (T_m, melting temperature). Addition of ligand to the sample stabilizes protein and therefore causes shift in melting temperature.

Nano-DSF measurements were performed using a Prometheus NT.48 instrument (NanoTemper Technologies, Germany). Purified DgoT was diluted to final concentration of 0.1 mg/ml with SEC buffer (20 mM HEPES, 150 mM NaCl, 0.05% LMNG, pH 7.4) and supplemented with galactonate. For measurements at other pH, SEC buffer was exchanged using a PD-10 desalting column to the solution containing 150 mM NaCl, 0.05% LMNG to remove HEPES. One of the following components was added from 200 mM stock to final concentration of 40 mM to the sample to set desired pH: AMPSO pH 9.27, Tricine pH 8.3, HEPES pH 7.4, MES pH 6.2. Next, samples were loaded into NanoDSF Standard-grade capillaries and transferred into a Prometheus NT.48 NanoDSF device. Thermal unfolding was detected during heating in a linear thermal ramp (1 °C/min; 20–95 °C) with an excitation power of 60%. Changes in the emission wavelengths of tryptophan fluorescence at 330 nm and 350 nm were monitored. The ratio between the emission intensities at 350 and 330 nm (F_{350}/F_{330}) was used to track the structural changes with increasing temperature. Melting-point temperatures (T_m) were calculated using the peaks in the first derivative of the signal data.

2.3 Data analysis

Experimental data were analyzed using in-house python scripts. Fits were performed using functionality of SciPy library. All SSME experiments were performed on at least three different sensors, each condition was measured at least twice to ensure reproducibility.

Molecular dynamics simulations were analyzed with GROMACS-based tools and in-house python scripts using MDanalysis, NumPy and SciPy libraries. Simulation trajectories were visualized using VMD [76] or PyMOL software.

3 Results

In this chapter, we propose a model of coupled transport mediated by DgoT. In this work, the experiments were primarily guided by the results of MD simulations, therefore we present them first. Although the transport cycle scheme can be devised using MD simulations alone, we find it useful to mention a few experimental results early on. We determined experimentally that DgoT symports two protons together with one galactonate molecule (Figure 3.22), therefore our scheme must account for the transport stoichiometry. For R47Q DgoT, we demonstrated that galactonate binding does not induce conformational changes in the protein (Figure 3.25), hence we looked closely into its role in the substrate binding. D46 and E133 are the only two negatively charged residues in the transmembrane part of the protein (Figure 1.4), therefore they were chosen as main candidates for proton acceptors. With MD simulations, we investigated the roles of protonation and substrate in transport and energetics of the major conformational changes. Understanding of these aspects allowed us to reconstruct complete transport cycle of DgoT and determine individual roles of several essential residues in the protein. We backed these findings up with various experimental techniques that are described in more detail later in this chapter.

3.1 Computational studies of DgoT

3.1.1 *Apo* simulations reveal flexible extracellular gate

"We studied H^+ and galactonate binding to outward-facing DgoT using unbiased all-atom MD simulations. For this, we used the crystal structure of E133Q DgoT [43] (PDB ID: 6E9O) as the starting conformation after reverting the mutation. First, we studied the effect of protonation on *apo* protein dynamics. The crystallographic galactonate molecule was removed from the binding site, and four different systems were generated with D46 and E133 individually protonated or deprotonated. We observed reproducible movements of the TM1 and TM7 transmembrane helices, with most flexibility in the most extracellular segments (around residues 48–52 of TM1 and 271–275 of TM7). This behavior resembles

the proposed role of TM1 and TM7 as gating helices that mediate transitions between open and occluded conformations, as in other MFS members [9,23,77].

When D46 and E133 were unprotonated, the extracellular gate (defined as distance between the center of mass of the Cα atoms of residues 48–52 and residues 271–275) was mostly closed (Figure 3.1A, D). Protonation of both D46 and E133 locked the gate in an open conformation (Figure 3.1B, D), and protonation of only one of these residues only partially opened the gate (Figure 3.1D). These changes in position of the gating helices were coupled to local rearrangements in the N-terminal domain. Charged, but not protonated, D46 and E133 interact with R126 and R47, respectively. After protonation, the side chain of D46 moves away, allowing R126 to interact with N49. This results in rotation of TM1 and stabilization of the open state of the extracellular gate (compare Figure 3.1A, B). In addition, protonation of E133 releases R47, which then becomes available to interact with the carboxyl group of galactonate." [A]

Movement and bending of gating helices result in rearrangement of hydrophobic residues located in TM1, TM2, and TM7. In conformation with closed gate side chains of F72, F274, L275 and W277 are tightly packed together (Figure 3.2A), forming a neck region with little to no water molecules in it (Figure 3.2C, blue line). In contrast, in conformation with open gate side chains of phenylalanine and tryptophan residues are moved away from the central pore (Figure 3.2B), resulting in better hydration of the neck region (Figure 3.2C, green line).

Since in conformation with closed gate key residues are separated from the extracellular solution (Figure 3.2A), their protonation may present a challenge in this conformation. However, in simulations with deprotonated D46 and E133 extracellular gate exhibited open conformation in some replicas. Therefore, we presume that gate opening is a stochastic process rather than a result of protonation event. Instead, role of protonation is stabilization of conformation with open gate, which is required for substrate binding.
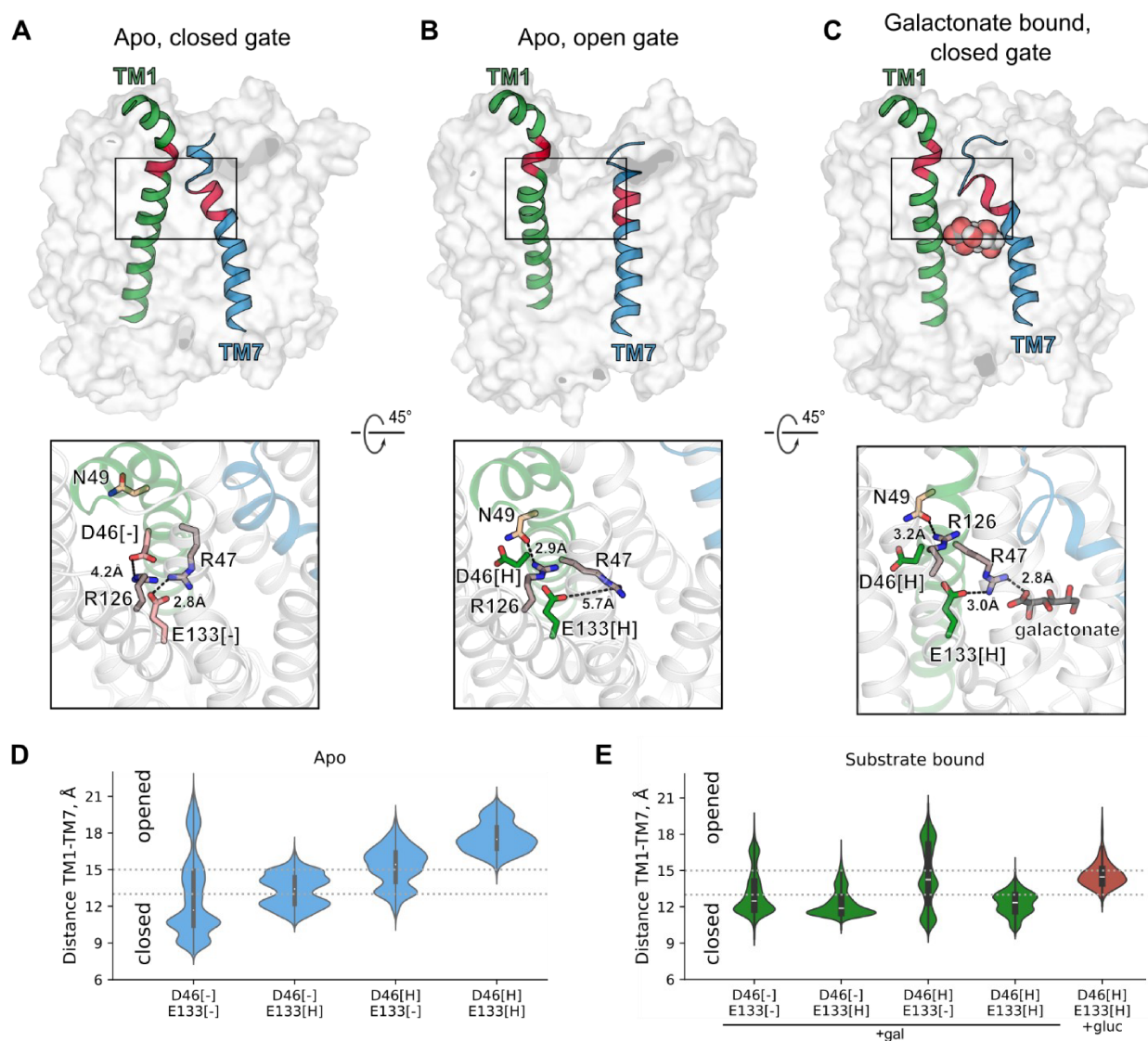
Figure 3.1. Extracellular gate dynamics in unbiased MD simulations. (A-C) Top row: snapshots showing the arrangement of the gating helices TM1 and TM7 in unbiased MD simulations with DgoT in outward-facing conformation with D46 and E133 deprotonated (A), protonated (B) and galactonate bound (C). The red colored parts of the helices correspond to residues 48-52 (TM1) and residues 271-275 (TM7). Bottom row: representative snapshots showing interactions around D46 and E133. (D-E) Probability densities for extracellular gate opening in apo (D) and substrate-bound (E) MD simulations with DgoT in the outward-facing conformation with different protonation states of D46 and E133. The TM1-TM7 distance is measured as distance between center of mass of Cα atoms of residues 48-52 and residues 271-275. The figure was adapted from [A].

"We performed unbiased simulations with the *apo* outward-facing structure in presence of 100 mM galactonate in solution. Before starting the MD simulations, D46 and E133 were protonated or deprotonated individually. For each system, at least one of the two key residues was protonated (five replicates were used). We observed spontaneous galactonate binding events in two simulations in which both D46 and E133 were
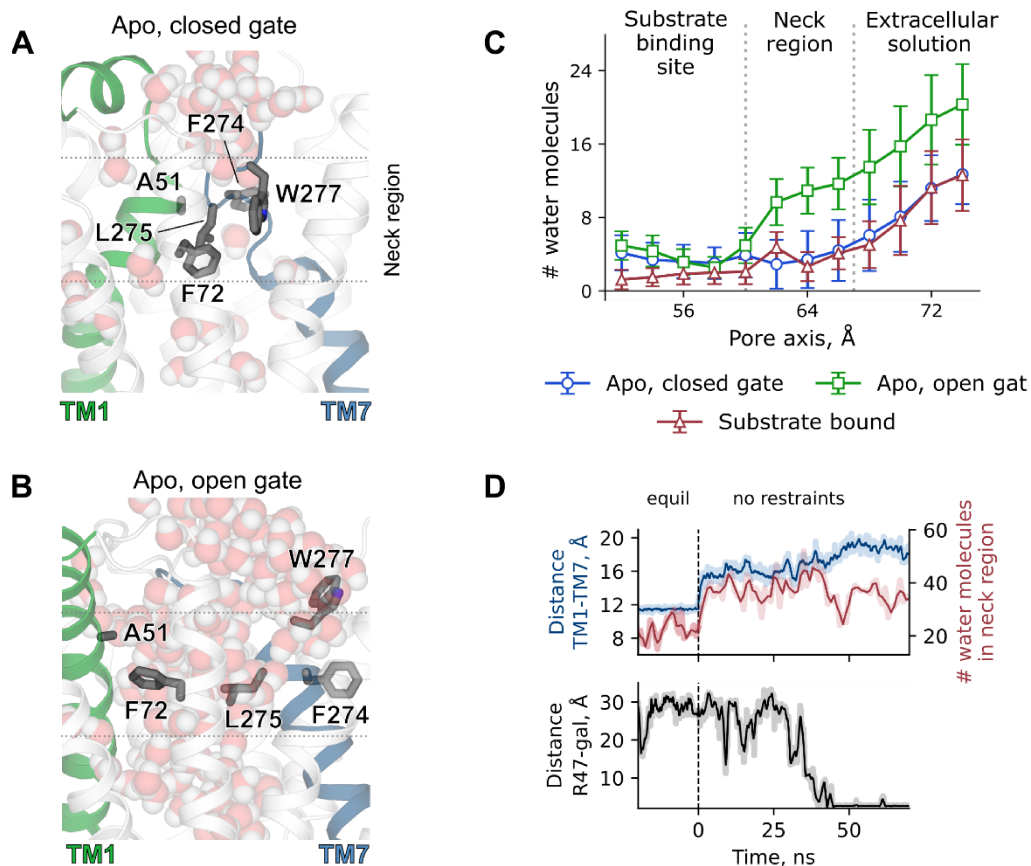
Figure 3.2. Hydrophobic interactions at the level of the extracellular gate. "(A, B) Snapshots showing water molecules distribution in simulations with DgoT in outward-facing conformation with closed (A) and open (B) extracellular gate. Representative snapshots were taken from unbiased MD simulations with both D46 and E133 either deprotonated (A) or protonated (B). (C) Hydration profile of the pore represented by the number of water molecules in 2 Å sections along the pore axis. Trajectories with following parameters were used for analysis: apo, closed gate – D46 and E133 protonated; apo, open gate and substrate bound – D46 and E133 protonated. In apo simulations only frames with minimum distance between side chains of F72 and W277 < 12.5 Å (blue line) or > 12.5 Å (green line) were used. (D) Time course of extracellular gate opening (measured as TM1-TM7 distance, as in Figure 3.1), number of water molecules in 10 Å section near extracellular gate (z coordinate between 58 and 68 Å) and galactonate binding to DgoT (measured as minimum distance between galactonate molecule and guanidinium group of R47) in a trajectory with both D46 and E133 protonated. Shaded lines represent raw data from the trajectory, solid lines are moving averages." [A}

protonated and in one simulation in which only E133 was protonated. In all cases, opening of the extracellular gate preceded the entry of galactonate into the binding site (Figure 3.2D)." [A]

3.1.2 Galactonate binding induces closure of the extracellular gate

According to "clamp-and-switch" model, substrate binding to MFS transporter induces conformational transition of the protein into oppositely open conformation [2]. As a first step of this major structural rearrangement, gating helices play an important role by bringing the transporter from open to occluded conformation.

"Unbiased simulations of outward-facing DgoT with bound galactonate revealed the reduced flexibility of gating helices TM1 and TM7 (Figure 3.1C) and the extracellular gate is more likely to be in a closed conformation (Figure 3.1E). Although the gate could close regardless of the protonation state of key residues in substrate-bound DgoT, the equilibrium was shifted toward the conformation with closed extracellular gate in simulations with only E133 protonated or with both D46 and E133 protonated (Figure 3.1E).

Protonation of E133 permits the R47 side chain to simultaneously interact with the carboxyl groups of galactonate and E133, resulting in a more compact arrangement of the charged residues (Figure 3.1C). Substrate interaction with residues on both gating helices facilitates bending of the TM7 extracellular segment toward TM1." [A]

If instead of galactonate molecule its epimer gluconate was present in the binding site, extracellular gate remained partially open (Figure 3.1E). This result further indicates that conformational change takes place after substrate recognition and cannot be induced by a molecule other than galactonate.

3.1.3 Substrate release from inward-facing DgoT

"We next ran unbiased simulations with the inward-facing structure of DgoT after placing a galactonate molecule into the binding site. The position for the substrate molecule

was determined by aligning the crystal structures of the *apo* protein in the inward-facing state (PDB ID: 6E9N) with the galactonate-bound outward-facing state (PDB ID: 6E9O). Galactonate in solution is predicted to have a pKa of 3.9 and, therefore, should bind to outward-facing DgoT in its deprotonated form. However, we also considered the possibility that galactonate could be protonated in the binding pocket due to changes in the local environment and then released in its protonated form. In 7 out of 38 individual 500 ns unbiased MD simulations, we observed spontaneous galactonate release (Table 3.1).

Protonation state: deprotonated (-) or protonated (H)			Total number of replica (500 ns long each)	Number of replica where substrate release was observed
D46	E133	galactonate		
-	-	-	4	2
-	H	-	10	3
H	-	-	10	0
H	H	-	5	0
-	H	H	4	2
H	-	H	4	0

Table 3.1. Substrate release in unbiased MD simulations with inward-facing DgoT with galactonate bound.

In all cases, the substrate passed between TM4 and TM10, which act as gating helices on the intracellular side of the protein. The intracellular segments of these helices (around residues 139–143 and 372–375, respectively) showed a high degree of flexibility in these simulations. The side chains of F137 and W373 formed a hydrophobic lock similar to the one at the extracellular gate (Figure 3.3A), and galactonate could only pass through this neck region if a water-filled pore was formed between the side chains of F137 and W373 (Figure 3.3B, C).

All spontaneous release events were observed in systems containing deprotonated D46, which suggests that the proton must be released from this residue prior to galactonate unbinding. Figure 3.3D shows probability densities for intracellular gate opening (measured as the minimum distance between the side chains of F137 and W373) for various protonation and substrate-binding states. The intracellular gate mostly assumed a closed conformation in the substrate-bound double protonated system. Deprotonation of D46 favored the open state of the gate in systems with a protonated or deprotonated substrate. When galactonate is protonated, the intracellular gate has limited flexibility and

becomes locked in the open state, when only E133 is protonated, and in the closed state, when only D46 is protonated. After release of the substrate and both H⁺ ions, the intracellular gate closes, permitting reorientation of the *apo* transporter to the outward conformation. We conclude that D46 deprotonation increases the probability of galactonate release from the inward-facing DgoT by promoting the opening of the intracellular gate." [A]
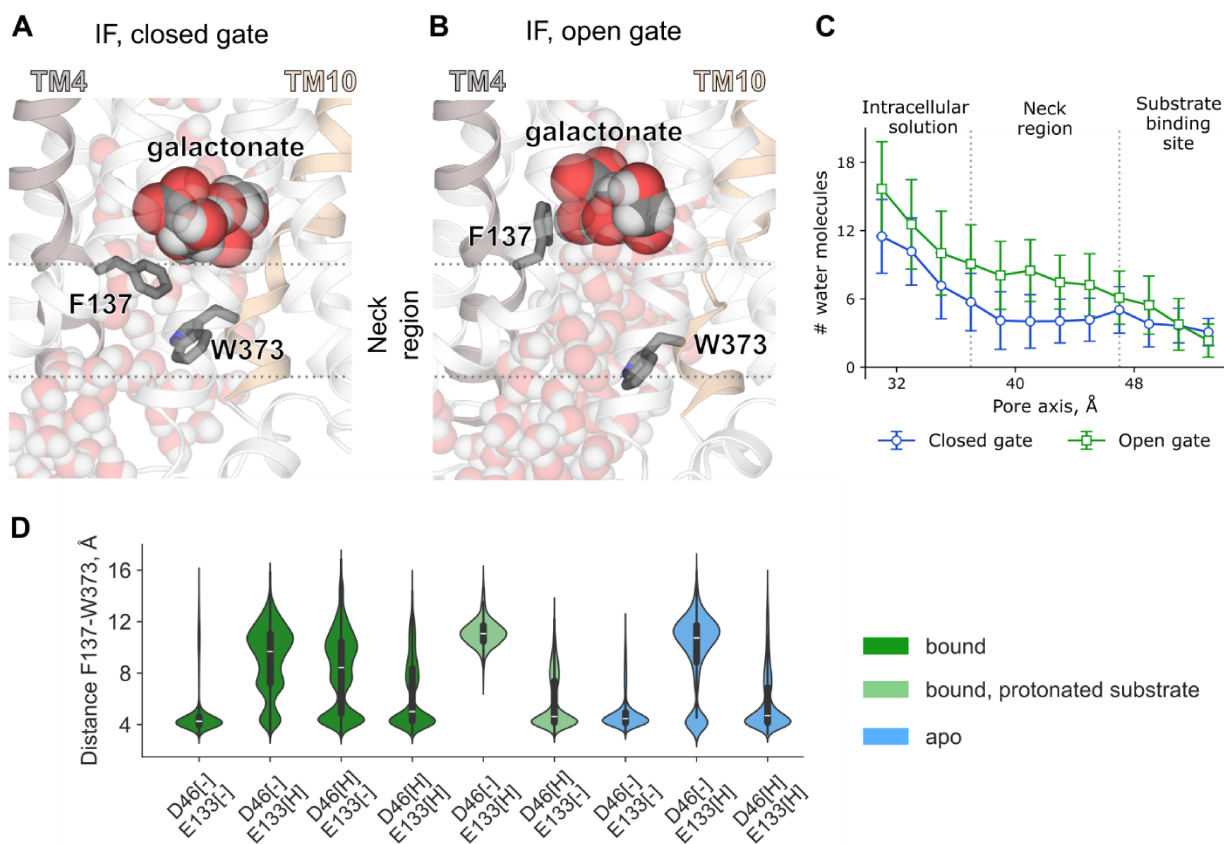
Figure 3.3. A hydrophobic lock within the intracellular gate regulates access to the binding site. "(A, B) Snapshots showing water molecule distribution in simulations with DgoT in inward-facing conformation with closed (A) and open (B) intracellular gate. Representative snapshots were taken from unbiased MD simulations with D46 and E133 protonated (A) or only E133 protonated (B). (C) Hydration profile of the pore, represented by the number of water molecules in 2-Å sections along the pore axis. Trajectories with following parameters were used for analysis: closed gate – D46 and E133 protonated; open gate – D46 deprotonated, E133 protonated. Substrate was bound to the protein in both cases. Only frames with minimum distance between side chains of F137 and W373 < 8 Å (blue line) or > 8 Å (green line) were used. (D) Probability densities for intracellular gate opening (measured as distance between side chains of F137 and W373) in MD simulations with DgoT in inward-facing conformation with different protonation states of D46 and E133 and galactonate and various substrate occupancies." [A]

Interestingly, the distance between backbone atoms of TM4 and TM10 (measured as distance between center of mass of Ca atoms of residues 139-143 and residues 372-375) was not evidently dependent on protonation of D46 and E133 (Figure 3.4). In simulations with protonated galactonate bound, distance between F137 and W373 changed significantly depending on whether D46 or E133 was protonated (Figure 3.3D, light green), while distance between TM4 and TM10 was mostly below 17 Å (Figure 3.4, light green). In *apo* simulations, distributions of TM4-TM10 distances are very similar for all tested combinations of D46 and E133 protonation (Figure 3.3D, blue), however, side chains of F137 and W373 were more frequently far apart in simulations with only E133 protonated (Figure 3.4, blue). This observation indicates that roles of extracellular and intracellular gates are different in DgoT. From extracellular side, highly flexible TM7 bends to bring together several large hydrophobic residues (F72, F274, L275, W277) and separate binding site from the bulk after substrate association. Intracellular parts of TM4 and TM10 are not so mobile, instead, movement of two hydrophobic side chains (F137 and W373) regulate access to the binding site.
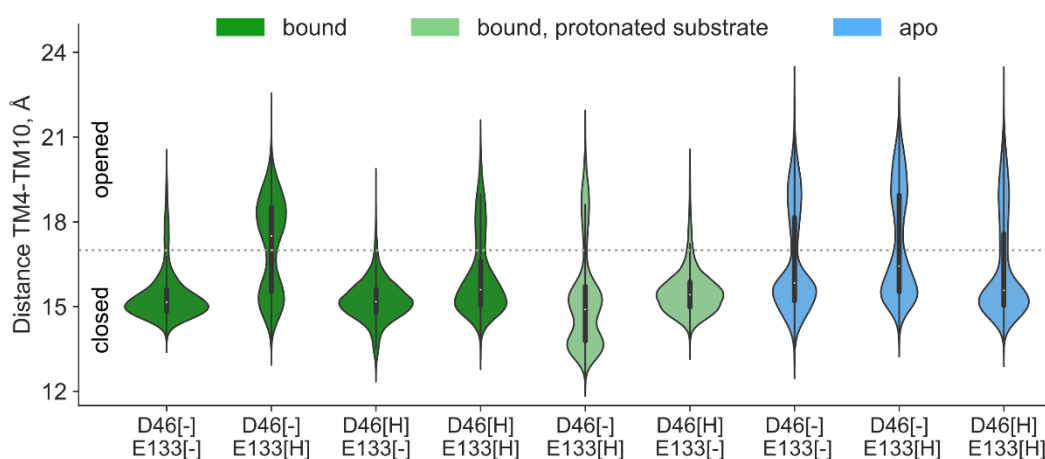
Figure 3.4. Distance between gating helices TM4 and TM10 in unbiased simulations. Distance was measured as distance between center of mass of Ca atoms of residues 139-143 and residues 372-375.

3.1.4 Proton release in inward-facing DgoT

Although in classical MD simulations, proton transfer events cannot be studied [49], some insights can be obtained from independent simulations with different protonation states of titratable residues. Since galactonate dissociated from inward-facing DgoT only in simulations in which D46 was deprotonated (Table 3.1), this residue is expected to deprotonate prior to substrate release. "However, in unbiased simulations with a system with two protons bound to the protein, the carboxyl group of protonated D46 is sequestered, making direct deprotonation to the intracellular solution unlikely (Figure 3.5A)." [A] To resolve this contradiction, we propose an alternative multi-step mechanism for proton release and support it with analysis of electrostatics in unbiased simulations with various protonation states of D46 and E133.
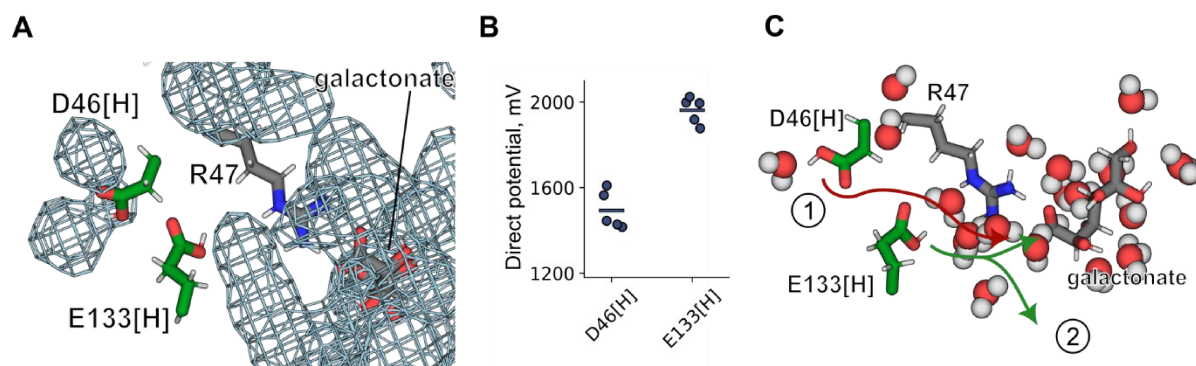
Figure 3.5. Proton release from D46. "(A) Water occupancy map in unbiased simulations with D46 and E133 protonated, contoured at an occupancy level of 0.1. (B) Electrostatic potential at the carboxyl groups of protonated D46 and E133 in the unbiased MD simulations." [A] Each data point is an average potential of a single trajectory. (C) Snapshot showing typical distribution of water molecules near D46, E133 and galactonate in unbiased simulations with inward-facing DgoT with D46 and E133 deprotonated. Hypothetical pathways for proton transfer from D46 (1) or E133 (2) are shown with arrows.

Electrostatic potential was quantified using *g_elpot* [66]. For calculation of time-resolved water-based potential, 0.3-nm sphere around carboxyl group the respective residue was chosen (CHARMM naming convention: OD1 and OD2 of D46, OE1 and OE2 of E133). Dehydrated frames (i.e. with no water molecules within selected sphere) were excluded from analysis. For calculation of the electrostatic potential sensed by individual

residues, 0.15-nm sphere around the following atom groups was considered (CHARMM naming convention: OD1, OD2, and CG of D46 and OE1, OE2, and CD of E133. Since protonated D46 and E133 are chemically identical, comparison of direct potential values is justified [67].

In system with D46 and E133 protonated, the electrostatic potential was higher on the carboxyl group of glutamate (Figure 3.5B), indicating that E133 is likely the first residue to release its proton. Distribution of water molecules near protonation sites suggests that proton transfer from E133 is more plausible than from D46 (Figure 3.5C). To preserve coupling, proton must be either transferred to carboxyl group of the substrate or released
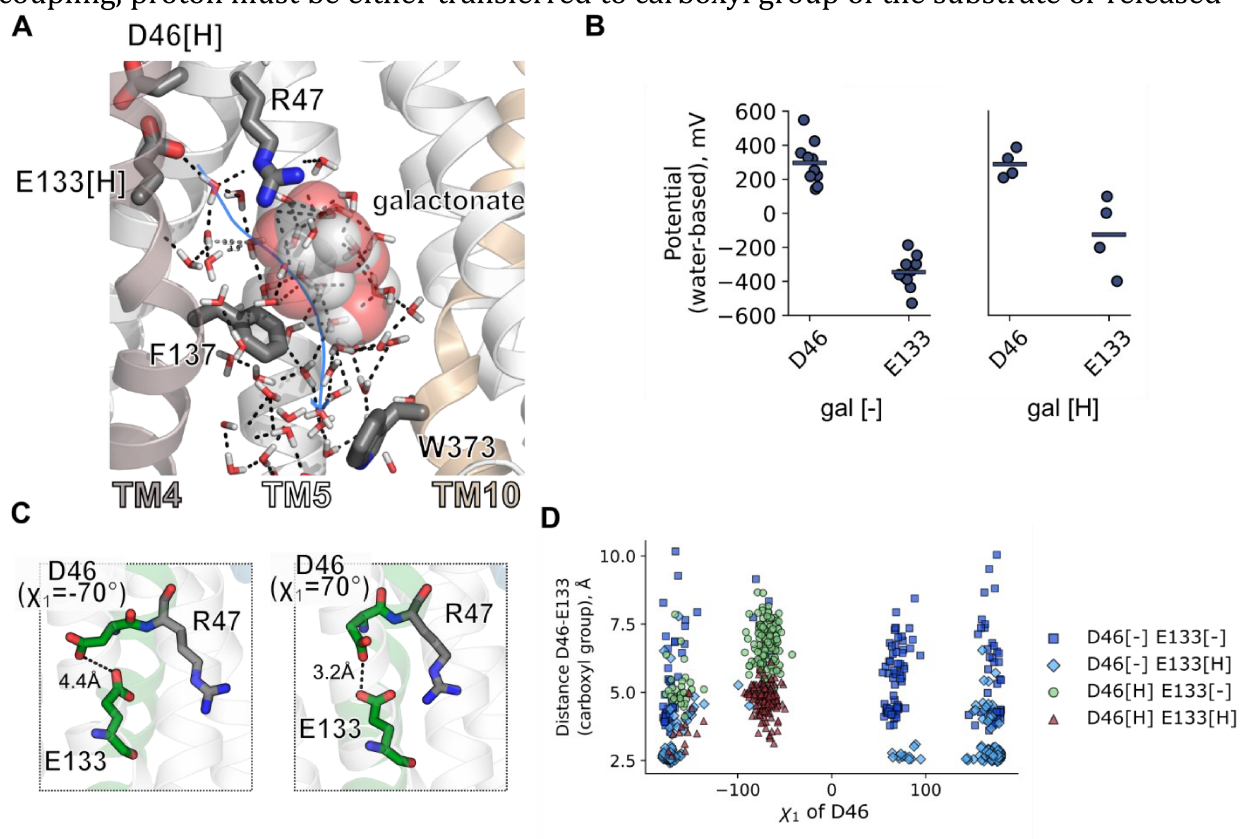
Figure 3.6. Intermediate steps in proton release mechanism. (A) Snapshot from a trajectory with substrate-bound inward-facing DgoT with D46 and E133 protonated showing a plausible water-mediated pathway for proton release from E133 to intracellular solution (blue arrow). (B) Electrostatic potential in 0.3-nm hydrated sphere around the carboxyl group of protonated D46 and deprotonated E133 in the unbiased MD simulations. Each data point is an average potential of a single trajectory. (C) Snapshots showing different configurations of D46 side chain in unbiased simulations with substrate-bound inward-facing DgoT with both D46 and E133 protonated (*left*) or only E133 protonated (*right*). (D) Relation between the first side chain dihedral angle of D46 (χ_1) and minimum distance between carboxyl groups of D46 and E133 in unbiased simulations with substrate-bound inward-facing DgoT.

to intracellular solution. In the first case, potential pathway for proton transfer might be mediated by several water molecules between the protonated E133 and galactonate (Figure 3.5C). In the latter case, proton release might seem more challenging, since proton must pass closed hydrophobic gate (Figure 3.3D) on its way to intracellular solution. However, a potential water-mediated proton pathway to the bulk still can be found near the hydrophobic region, closer to TM5 (Figure 3.6A).

In both scenarios, deprotonation of E133 changes local environment of protonated D46. Water-molecule electrostatic potential near the side chain of E133 becomes substantially lower than that near D46 side chain, promoting the possibility of proton transfer from D46 to E133. Although in our simulations side chain of protonated D46 is usually turned away from E133 (Figure 3.6C *left*), in simulations with different protonation states it could easily move towards E133 (Figure 3.6C *right*). The latter configuration is presumably more favorable for proton transfer, and to achieve it, D46 must flip its side chain towards E133 (Figure 3.6D).

After proton transfer from D46 to E133 intracellular gate opens (Figure 3.3D) and galactonate can be released, either protonated or deprotonated (Table 3.1). In *apo* system with protonated E133, gate stays open (Figure 3.3D), making final deprotonation of E133 to intracellular solution via water molecules network plausible.

3.1.5 Markov state modeling

Available structures of related MFS transporters provide important insights about different intermediate conformations along inward-outward isomerization [43,44,78–81]. For DgoT, crystal structures capture protein in inward open (PDB ID: 6E9N) and outward occluded (PDB ID: 6E9O) states. During unbiased simulations protein conformation showed some flexibility depending on protonation state and ligand occupation. When substrate was bound, DgoT was predominantly in inward- or outward-occluded conformation (depending on crystal structure used as starting conformation), while *apo* protein was more dynamic and displayed more open conformation depending on protonation state of D64 and E133 (Figure 3.1). However, protein conformation remained

close to that in the corresponding starting state and we did not observe transitions between inward and outward facing conformations in single trajectories. To describe energetics of such slow process we employed Markov state modeling that can stitch together information about conformational transitions from multiple short trajectories.

First, suitable collective variable that describes conformational state must be chosen. "We featurized the trajectory data using a set of interdomain Cα atomic distances and applied time-lagged independent component analysis (tICA) to find the slowest components in the dataset. The first tICA-eigenvector discriminates between the inward- and outward-facing conformations (Figure 3.7A), and the second correlates with the (open or closed) state of the extracellular gate (Figure 3.7B, Figure 3.8A). Two systems were
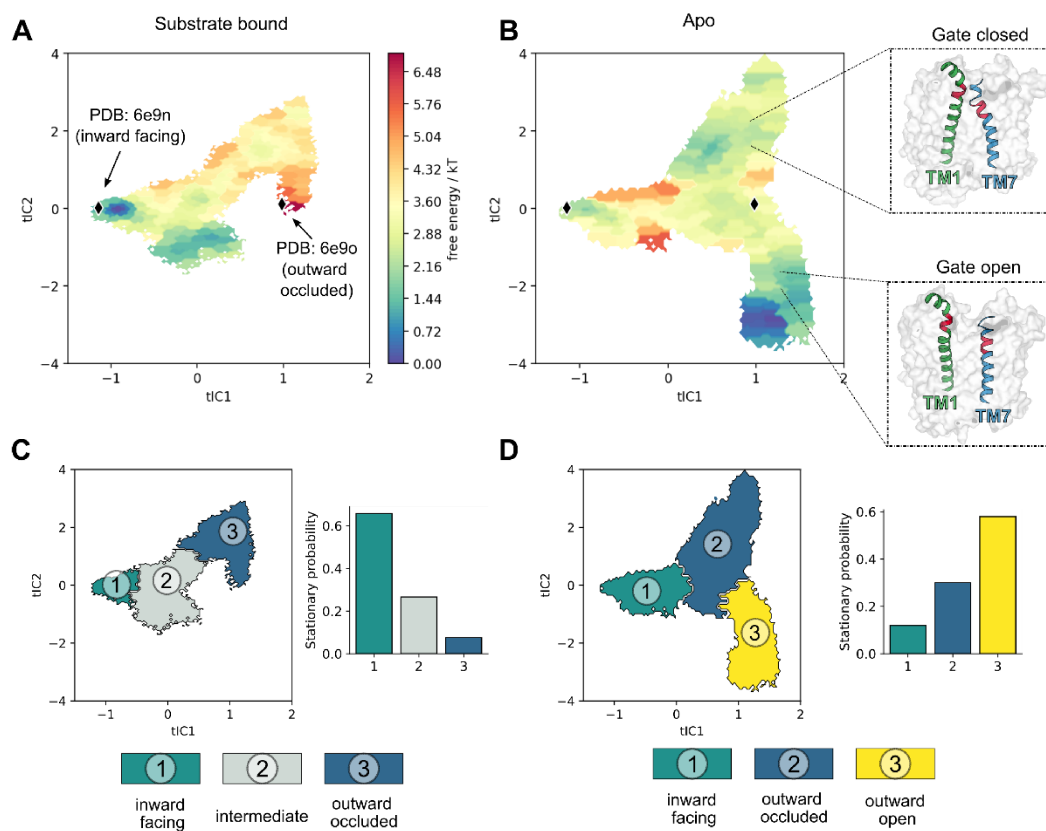
Figure 3.7. Major conformational changes described with MSM. "(A, B) Free energy landscape for DgoT in substrate-bound (A) or apo (B) state. Protein conformations captured in DgoT crystal structures were projected onto the tICA space (black points). Representative snapshots illustrate how the value of tIC2 correlates with the degree of opening of the extracellular gate. (C) Left: coarse representation of intermediate metastable states obtained with PCCS for substrate-bound DgoT. Right: Stationaty probabilities of such metastables states. (D) Same representation as in (C), for apo DgoT." [A]

chosen as most relevant for transport: (i) galactonate-bound DgoT with protonated D46 and E133 (responsible for translocation of the substrates across the membrane) and (ii) *apo* DgoT with deprotonated D46 and E133 (responsible for reorientation of the empty transporter after substrate release). For both systems, we sampled all relevant intermediate conformations with a total of ~26 μs (substrate-bound) and ~35 μs (*apo*) unbiased MD simulations, and constructed Markov state models (Figure 3.7A, B). The free energy surfaces for both systems revealed a high energy barrier separating the inward- and outward-facing conformations. In the presence of galactonate, the barrier was lowered and the inward-facing state was the most energetically favorable. In contrast, the system with the *apo* protein preferred outward-facing conformations such that DgoT favors inward galactonate transport. Moreover, *apo* DgoT easily switches between outward-occluded and outward-open states, whereas substrate-bound DgoT adopts only occluded conformations." [A]

"Perron-cluster cluster analysis (PCCA)[65], used to identify metastable states, revealed a difference in the direction of conformational changes between the *apo* and galactonate-bound states (Figure 3.7C, D), as previously observed in the free energy surfaces (Figure 3.7A, B). For substrate-bound DgoT, metastable states correspond to inward-facing, intermediate occluded, and outward-occluded states, with the highest probability for the inward conformation (Figure 3.7C). For *apo* DgoT, PCCA identified inward-facing, outward-
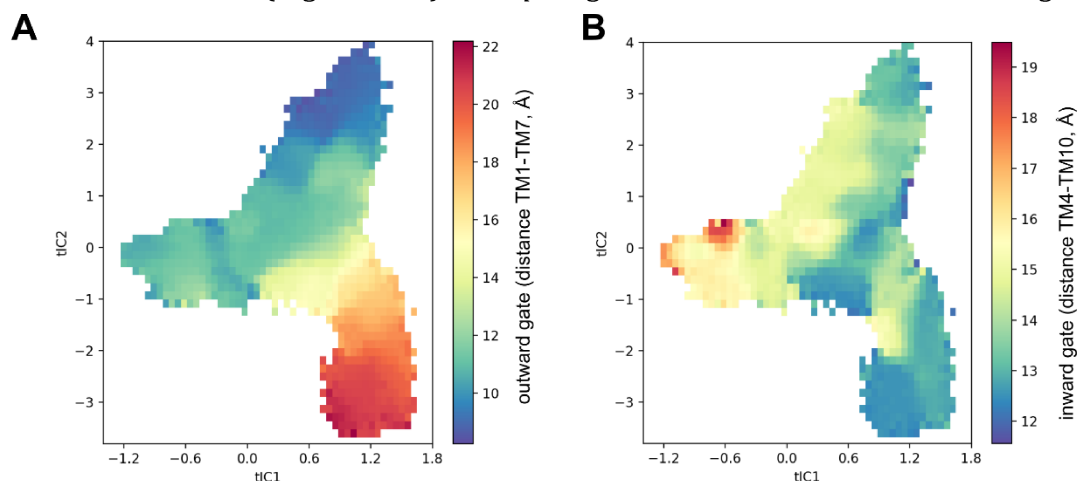
Figure 3.8. Correlation between tICA eigenvectors and distances between gating helices for *apo* system. Distances between gating helices measured as distance between center of mass of Cα atoms of residues 48-52 and residues 271-275 (A) or center of mass of Cα atoms of residues 139-143 and residues 372-375 (B).

occluded, and outward-open states (lower to higher probability). The outward-open state was not adopted by substrate-bound DgoT: it was only adopted by *apo* DgoT, for which it represents the state with highest stationary probability (Figure 3.7D). Given that substrate binding and release are fast processes compared with the conformational changes, our data show that reorientation of the empty transporter upon substrate release is the rate-limiting step in the transport cycle." [A]

For model validation, we performed a Chapman–Kolmogorov test with three metastable states for both the substrate-bound and *apo* DgoT systems (Figure 3.9D, E). For each system, the model prediction and re-estimation are in good agreement up to 3 lag times used for MSM construction.
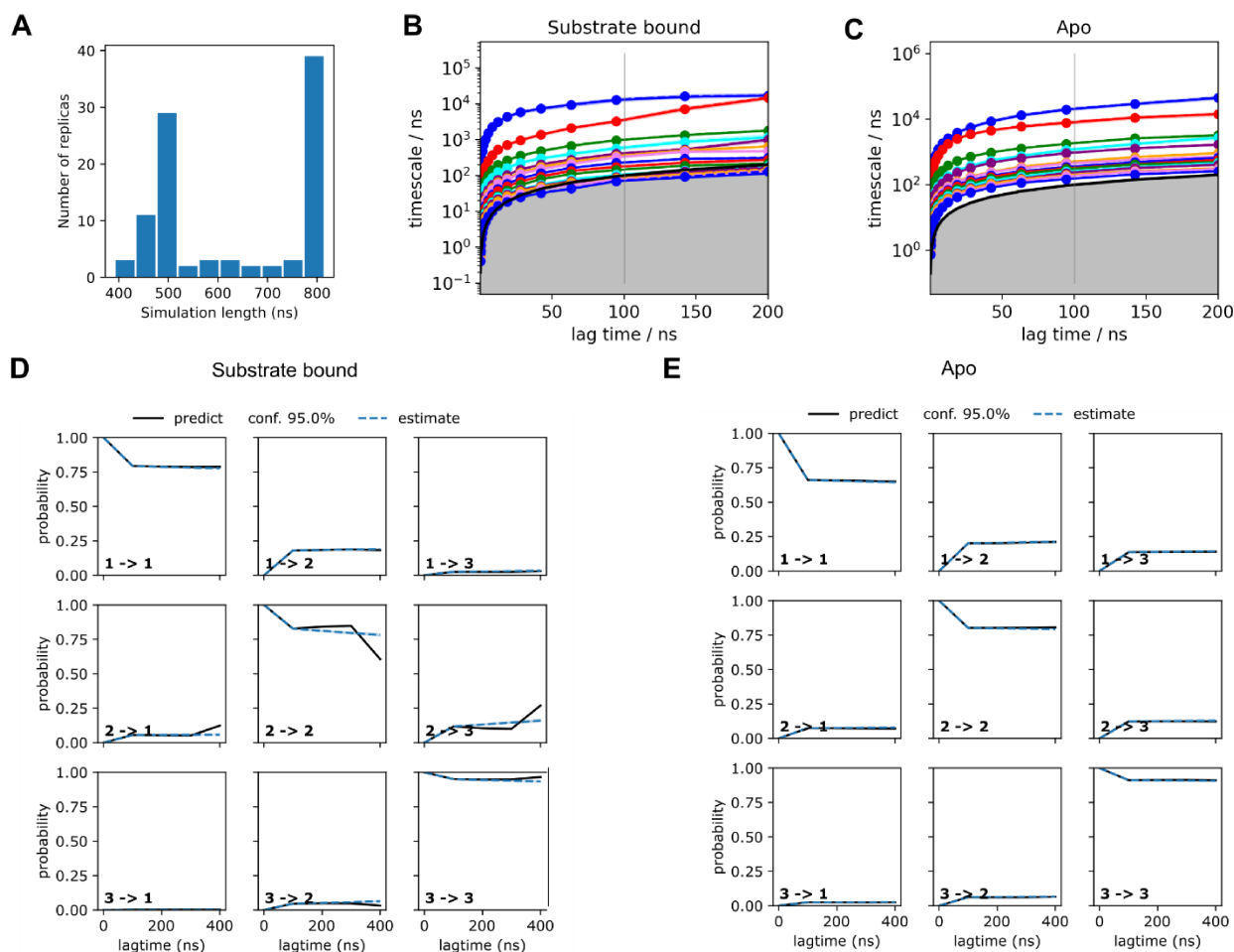
Figure 3.9. MSM validation. "(A) Histogram of lengths of trajectories used for MSM construction. (B, C) Implied timescales plot for systems with substrate bound (B) or apo (C) DgoT. Vertical line indicates lag time value chosen for MSM construction. (D, E) Chapman-Kolmogorov test for systems with substrate bound (D) and *apo* (E) DgoT." [A]

3.1.6 Role of R47Q in formation of occluded conformation

Our experimental data demonstrated distinct role for R47 in functional activity of the protein (Figure 3.25B). "The crystal structure of outward-facing DgoT [43], as well as our MD simulation data (Figure 3.1), indicates that R47 interacts with the carboxyl group of bound galactonate. To better understand the role of this arginine residue, we carried out MD simulations with the outward-facing DgoT structure in which R47 was mutated to glutamine and D46 and E133 were protonated. In apo simulations, both WT and mutant DgoT assumed similar conformations with an open extracellular gate (Figure 3.10A, *top*). However, galactonate failed to promote closure of the extracellular gate of R47Q DgoT (Figure 3.10A, *bottom*).
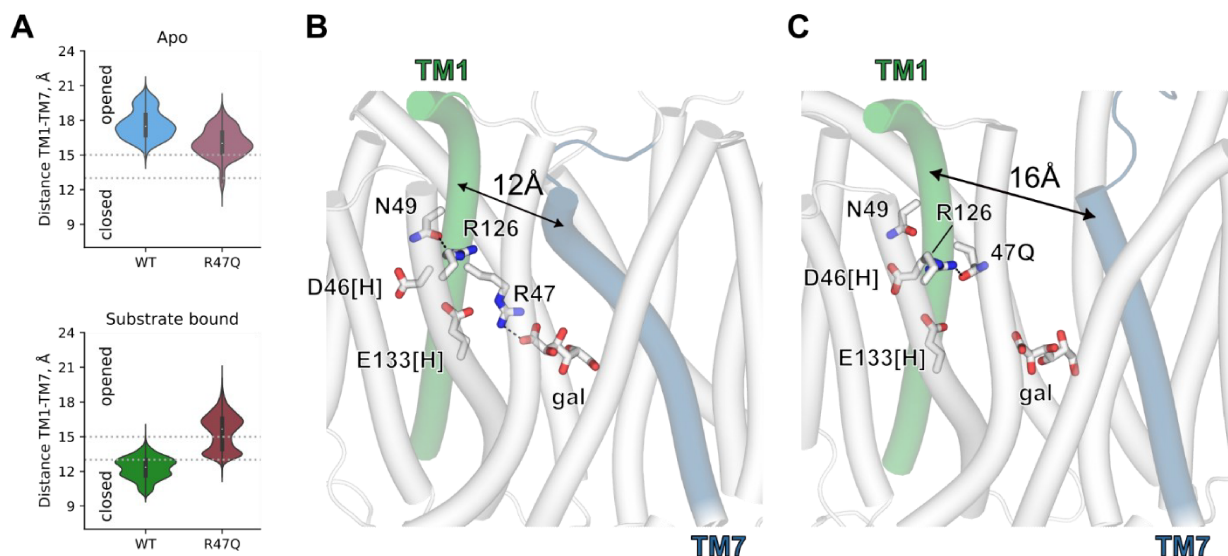
Figure 3.10. Extracellular gate dynamics in unbiased simulations with R47Q DgoT. (A) Probability densities for extracellular gate opening in apo (top) and galactonate-bound (bottom) simulations with WT and R47Q DgoT in outward-facing conformation with D46 and E133 protonated. (B, C) Snapshots showing the arrangement of the gating helices TM1 and TM7 in unbiased MD simulations with WT (B) and R47Q (C) DgoT in outward-facing conformation with D46 and E133 protonated and galactonate bound.

This difference is a consequence of changes in protein–substrate interactions. Galactonate is coordinated by multiple polar residues [82] in helices from both the N- and C-domains, with the substrate located closer to TM7 in simulations with the mutant than with the WT transporter (Figure 3.11A, B). Comparison of the substrate–protein

interactions revealed more stable contacts for galactonate with C-domain residues such as Q264, T372, and N393 in WT; residue 47 from the N-domain interacts with the substrate only in simulations with the WT protein (not the mutant protein) (Figure 3.11C)."
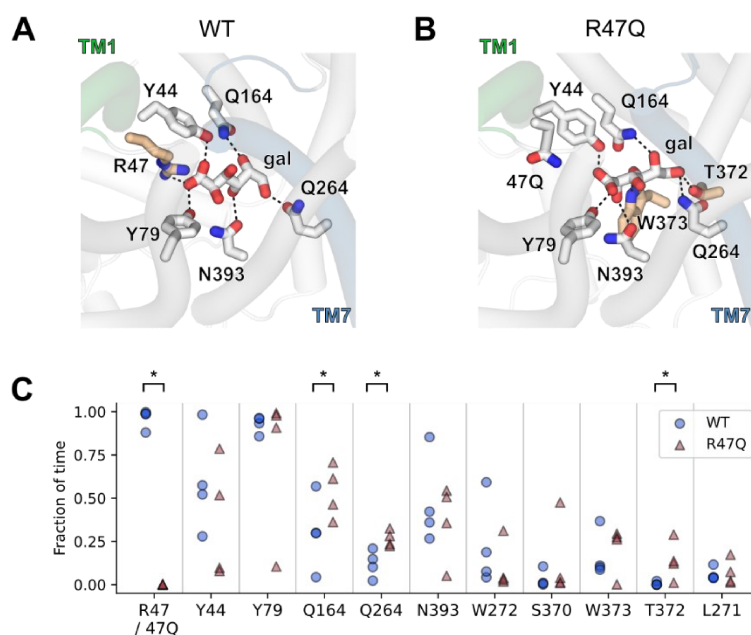
Figure 3.11. Changes in substrate-protein interaction in R47Q DgoT. "(A, B) Position of galactonate in the binding site of WT (A) and R47Q DgoT (B). (C) Fraction of frames, in which distance between the side chain of specified residue and closest oxygen atom of galactonate was equal or less than 2 Å. Each data point represents one of the four independent trajectories for WT (blue circles) and R47Q mutant (red triangles) DgoT. Significance was evaluated with the Mann–Whitney test, one-sided: *$p < 0.05$." [A]

3.1.7 Model of transport cycle of DgoT

Key elements of transport mechanism of DgoT were investigated with MD simulations and summarized in transport cycle model (Figure 3.12). Galactonate transport is achieved via alternating access mechanism, therefore isomerization between inward- and outward-facing conformations is required. Movement of the gating helices facilitated by protonation and substrate binding is a crucial part of major conformational changes. "In the outward-facing conformation, DgoT dynamically switches between states with a closed or open extracellular gate (states 1 and 2). Protonation of both D46 and E133 stabilizes the open-

gate conformation (state 2) and permits galactonate binding from the periplasmic side (state 3), followed by closure of the extracellular gate. Formation of the outward-occluded conformation is the initial step of major conformational change (state 4) that brings the protein into an inward-facing conformation (states 5, 6, and 7). In this conformation, deprotonation of D46 (state 6) results in an open intracellular gate that permits galactonate release (state 7). D46 deprotonation may occur via multiple pathways, as shown for proton transfer in other transporters [83]. Our multiscale simulations demonstrate that D46 deprotonation is possible via initial proton release from E133, either to the intracellular solvent or to galactonate, followed by H^+ transfer from D46 to E133 (state 6). Galactonate can unbind in a protonated or non-protonated form; protonation might promote galactonate release by weakening the electrostatic interaction with R47 and stabilizing intracellular gate opening. H^+ transfer from D46 to E133 (state 6) and subsequent release from E133 results in intracellular gate closure (state 7); reorientation of the empty transporter to outward-facing conformation completes the cycle (state 8)." [A]
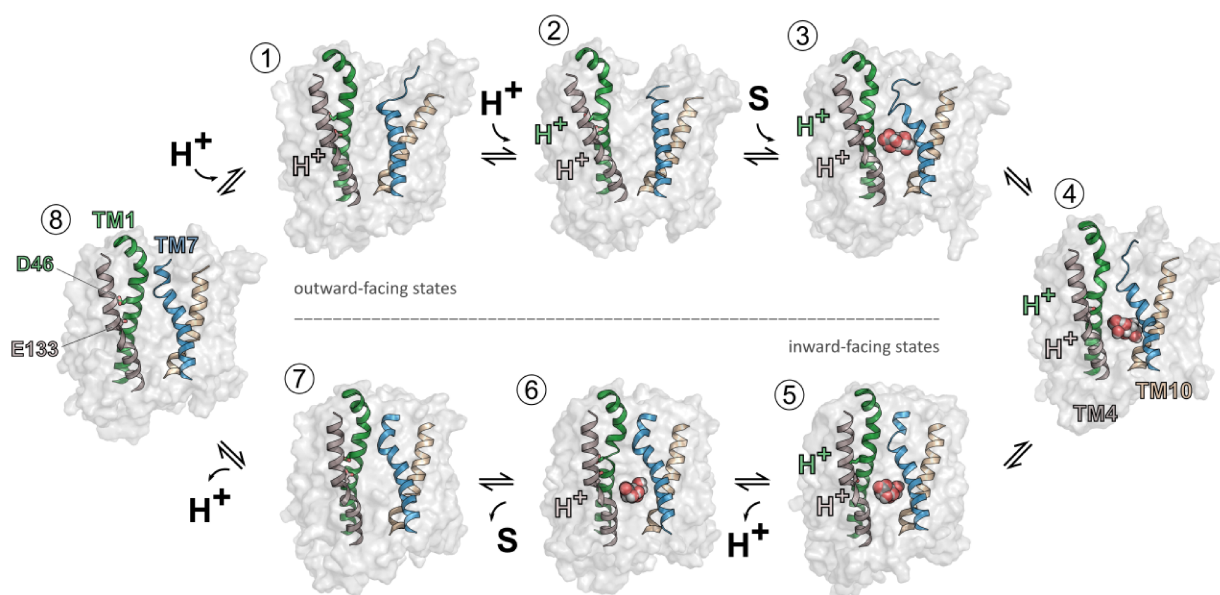
Figure 3.12. Model of DgoT transport cycle reconstructed with MD simulations.

3.2 Experimental characterization of DgoT transport

3.2.1 Analysis of expression, purification and reconstitution

Different DgoT variants were expressed in *E.coli* C41 cells and purified as described in methods. Size-exclusion chromatography profile demonstrated that purified DgoT behaved as a monodispersed protein. Elution peak was observed at 12.6 ml for WT protein and at 12.3-12.7 ml for mutant variants, which corresponds to monomeric protein in detergent micelles (Figure 3.13A). SEC fractions with absorbance no less than 1/3 of peak value were pooled and stored at -80 °C until further use. The quality of the purified protein was additionally checked by the SDS-polyacrylamide gel electrophoresis (SDS-PADE). Coomassie stain of purified DgoT separated by SDS-PAGE confirms that major elution peak along fractions 21-27 corresponds to monomeric protein with apparent molecular weight of approximately 35 kDa (Figure 3.13B).

Individual purification steps were analyzed by western blot, as illustrated on Figure 3.14. Typical yield of purified WT DgoT was 4.5 mg from 1L of cell culture.
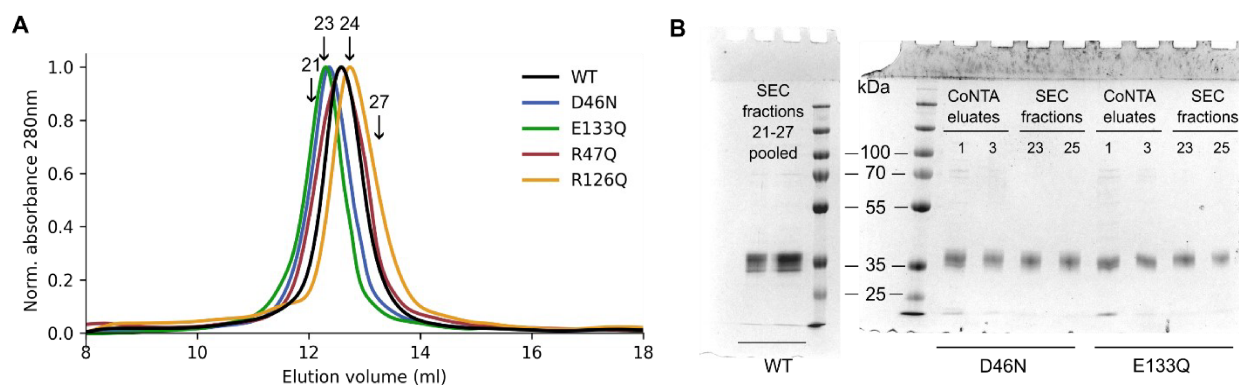
Figure 3.13. Purified DgoT variants. (A) Size-exclusion chromatography (Superdex 200Increase 10/300 GL) of purified WT DgoT and single mutants. Absorbance at 280 nm for each profile is normalized to the value of peak signal. Vertical arrows indicate fraction numbers that were subjected to SDS-PAGE analysis. (B) Coomassie stain of purified DgoT separated by SDS-PAGE.
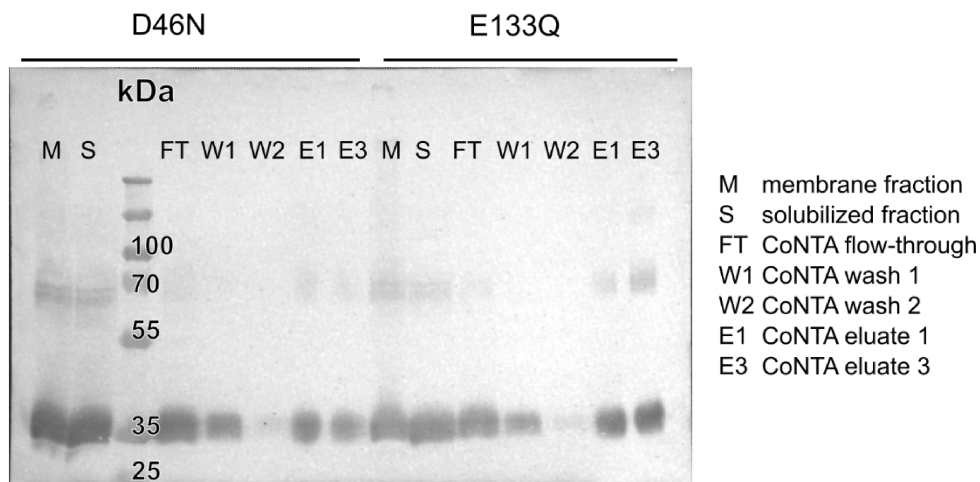
Figure 3.14. Purification stages. Western blot for D46N and E133Q mutants at different purification steps.

For SSME experiments, purified DgoT was reconstituted into proteoliposomes. Efficiency of reconstitution was evaluated by centrifugation on a 10-40% sucrose gradient. After centrifugation tube content was divided into 11 layers which were carefully collected from top of the tube. Lipids were identified visually in fraction 7 from top as corresponding fraction was significantly more turbid than others. Protein amount in separated fractions was estimated by SDS-PAGE. Most protein was found to be in fraction 7 and smaller amounts were also found in fractions 6 and 8, likely due to unavoidable partial mixing of adjacent fractions. Protein was found neither in the lowest fraction, where aggregates might be expected, nor in upper sucrose layers, where poorly incorporated solubilized protein would be expected to appear (Figure 3.15).

Presence of protein in the same fraction as lipids indicates successful insertion into liposomes. Although sometimes such a result could be false positive (for example, if nonincorporated proteins stick onto the surface of liposome membranes instead of being inserted) [84], this assay can be used as evidence of good quality of reconstitution procedure.
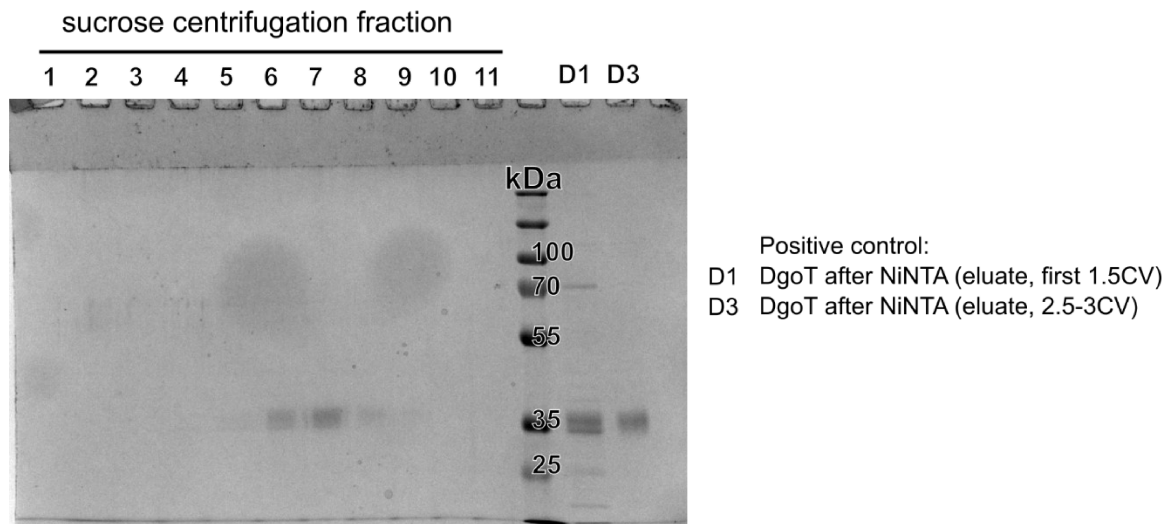
Figure 3.15. Evaluation of reconstitution efficiency. Sucrose gradient fractions analyzed with SDS-PAGE (lanes 1-11, sucrose concentration is highest in the last fraction). Partially purified protein in detergent is added for reference (lanes D1 and D3).

3.2.2 Transport assay in bacterial cells

We demonstrated that DgoT is functional when overexpressed in *E.coli* cells using pH electrode-based assay. *E.coli* C41 cells were grown in 50 ml TB medium, protein expression was induced by IPTG and continued for 2 or 3 hours. Then the cells were washed in non-buffered assay solution and pH of bacterial suspension was continuously measured. Addition of galactonate resulted in immediate increase in pH, although galactonate solution had slightly lower pH than the bacterial suspension. Galactonate epimer gluconate, as well as cells without DgoT, showed no changes in pH upon addition of substrate, confirming that observed pH increase reflect transport activity of DgoT.

In each experiment maximum value of pH was reached in less than a minute after the addition of galactonate. We compared maximum increases of pH observed in experiments with different mutants to identify amino acids that are important for transport activity.
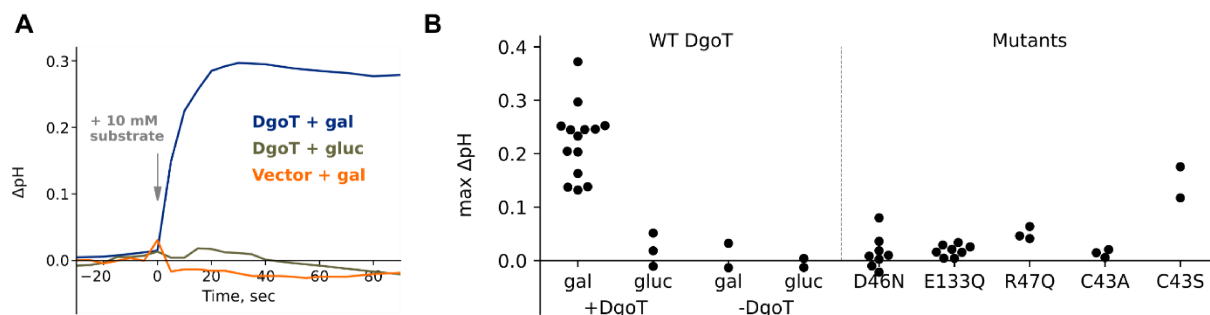
Figure 3.16. pH electrode-based assay. "(A) pH changes induced by H$^+$-coupled galactonate, but not gluconate transport into *E.coli* cells expressing WT DgoT or transformed with the same vector without the DgoT gene. (B) Maximum pH changes observed in experiments with *E.coli* cells expressing different DgoT variants." [A]

Since point mutations can cause significant changes in protein expression, thus leading to misinterpretation of results of pH electrode-based assay, we compared expression levels of different DgoT variants in *E.coli* cells used in this experiment using western blot (Figure 3.17). Expression levels of all tested mutants were comparable with those of WT protein, confirming that cells expressing mutated DgoT variants respond differently to addition of galactonate due to changes in functional activity of the protein.
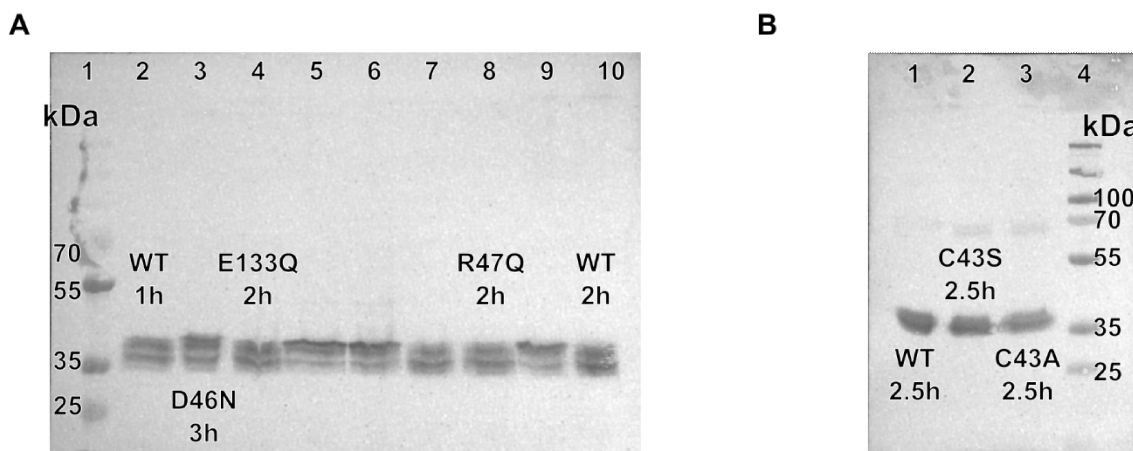
Figure 3.17. Expression of WT and mutated DgoT in *E.coli* C41 cells used for pH electrode-based assay. (A) Western blot for WT (lanes 2 and 10), D46N (lane 3), E133Q (lane 4) and R47Q (lane 8) mutants after 1-3 hours expression. (B) Western blot for WT (lane 1), C43S (lane 2) and C43A (lane3) after 2.5 hours expression. (A and B) All probes normalized to cell optical density. Lanes with WT DgoT refer to three independent protein expressions.

In comparison with SSME, this assay cannot provide information about kinetics of transport, however, can potentially detect electroneutral transport since it reports on changes in proton concentration in bacterial suspension. Since neither D46N nor E133Q mutants demonstrated significant pH change in response to galactonate addition, we concluded that neither of them can function as 1:1 H^+:galactonate symporter.

3.2.3 Substrate binding to purified protein does not induce pH changes

MD simulations suggest that two protons bind to DgoT and stabilize open conformation of the extracellular gate prior to galactonate (Figure 3.1). However, unbiased simulations alone were not conclusive the binding order and did not exclude alternative binding combinations, in which one or two protons bind after the substrate. Spontaneous galactonate binding was observed in simulations with E133 protonated, while protonation state of D46 did not seem to play critical role (paragraph 3.1.1).

To test experimentally the binding order of protons and substrate, we monitored pH changes in solution with solubilized protein upon addition of galactonate. The procedure was analogous to the transport assay with bacterial cells described in paragraph 3.2.2. Purified DgoT in detergent micelles in unbuffered solution was equilibrated at RT and, if necessary, supplemented with NaOH or HCl until pH of the sample reached reasonably stable value around 7.5. If galactonate binding results in protonation/deprotonation of the
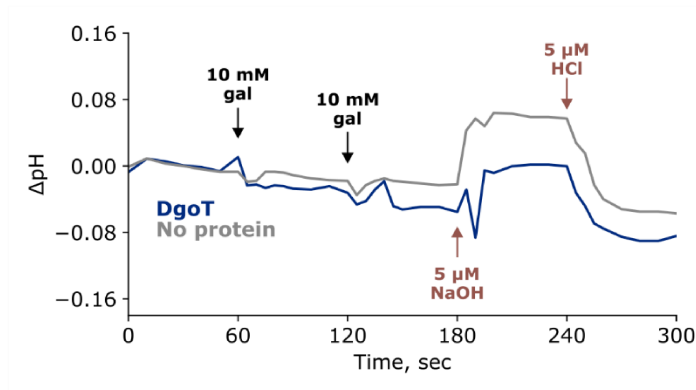
Figure 3.18. Changes in pH after addition of substrate to purified DgoT in detergent. Additions of 10 mM of galactonate are indicated by the black arrows.

protein, changes in proton concentration comparable with protein concentration (4 µM) upon substrate addition are expected [85]. "However, addition of 10 mM galactonate to purified DgoT in unbuffered detergent-containing solutions did not elicit any change in pH (Figure 3.18, blue line)." [A] To show that proton binding to the protein would have induced noticeable changes in pH of the solution, 5 µM of NaOH and HCl were added sequentially to the sample and resulted in changes in pH of 0.08-0.1 units. Measurements performed with purified DgoT and assay solution without protein were virtually undistinguishable (Figure 3.18, grey line). "This result suggests that the substrate binds to the protonated transporter and does not induce protonation of DgoT. Therefore, we conclude that DgoT binds two H^+ prior to galactonate association." [A]

Absence of galactonate-induced pH changes in DgoT-containing solution might be related to inability of solubilized protein to bind substrate. Therefore, galactonate binding to the purified protein was investigated with alternative techniques such as nanoDSF (Figure 3.27) and SSME (Figure 3.19B, Figure 3.24C). We conclude that purified protein in detergent micelles, as well as protein reconstituted into proteoliposomes, binds galactonate at millimolar concentrations, thus, our assay reports on the binding order of protons and substrate.

3.2.4 pH dependency of DgoT transport

DgoT transport was characterized with SSM-based electrophysiology, a method that relies on chemical substrate gradients to activate the transport activity of the protein. In single solution exchange experiments, transport was initiated by galactonate concentration jump, while pH of activating (A) and non-activating (NA) solutions was identical. All NA solutions were supplemented with non-substrate gluconate at the same concentration as galactonate in A solution to keep osmolarity of the two solutions identical and avoid artifacts caused by solution exchange.

50

At neutral and alkaline pH conditions, slowly decaying currents were recorded (Figure 3.19A). To test whether observed currents represent transport reaction, we compared current decays recorded with liposomes reconstituted at different lipid to protein ratios (LPR). At lower LPR each liposome has more proteins reconstituted, therefore membrane potential that counteracts chemical driving force and eventually leads to decay of transport is forming faster. Thus, for transport reaction systematical dependency of current decay on LPR is expected [86]. Slowly decaying currents in our experiments were notably influenced by LPR (Figure 3.19C), confirming that they represent continuous turnover.
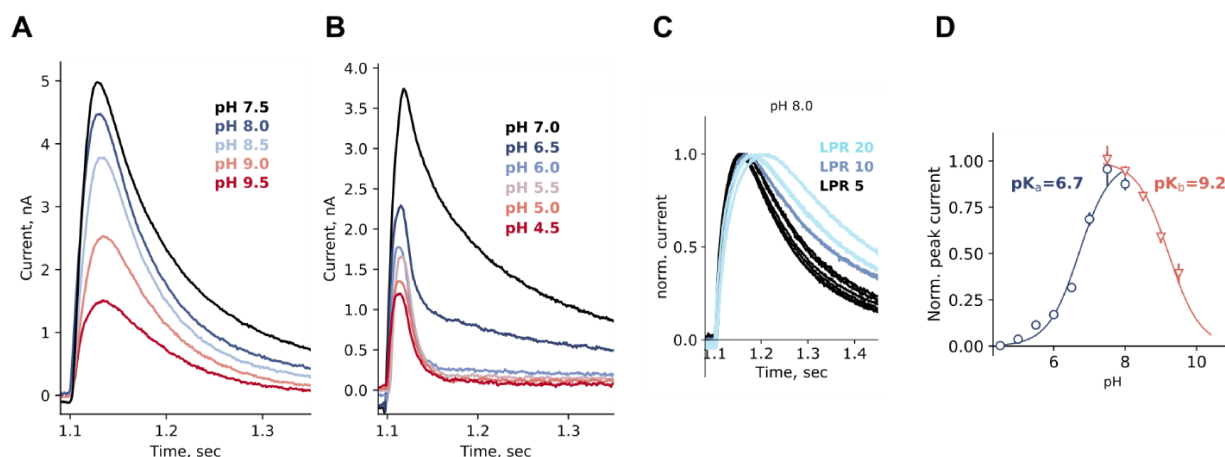
Figure 3.19. SSME currents under different pH conditions. (A, B) Alkaline (A) and acidic (B) deactivation of WT DgoT transport currents measured by SSME upon application of 10 mM D-galactonate concentration jump. (C) Transport currents at pH 8.0 from liposomes reconstituted with different LPR. Currents were normalized to their peak value for comparison of the current decay. (D) Peak currents and determination of apparent pK values. The solid lines are fits to the data using the equations described in Methods. The figure was adapted from [A].

We studied effect of pH on DgoT transport by performing single solution exchange experiments on the same sensor using solutions with various pH. Before each measurement, sensor was incubated with respective NA solution for several minutes to ensure equilibration of intraliposomal pH to the value of the new solution via passive proton flux across liposome membrane [87]. Efficiency of intraliposomal pH exchange was evaluated by comparison of traces recorded in duplicates (i.e. with the same solution pair). Solutions with pH from 9.5 to 4.5 in steps of 0.5 pH units were successively measured. Additionally, measurements with solutions with pH 7.0 were added at the beginning and at the end of the sequence as a rundown control to evaluate sensor stability during the whole experiment. Transport currents assumed a maximum amplitude at pH 7.5 and decreased at higher pH values (Figure 3.19A). Since galactonate transport is proton coupled, decrease in transport activity is expected under the conditions with lower proton concentration. Under acidic pH conditions, peak current amplitude was reduced and shape of current decay changed (Figure 3.19B). Slow transport component decreased as pH of the solution
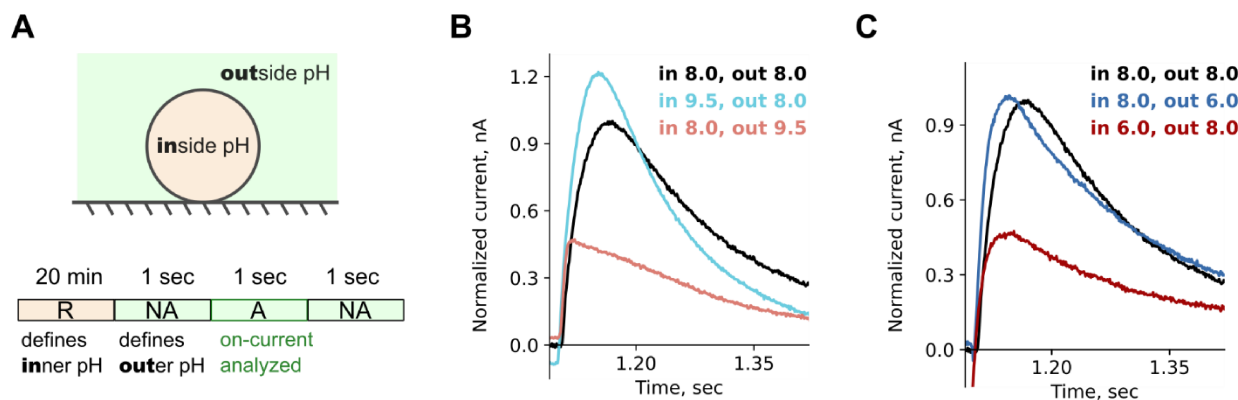
Figure 3.20. SSME experiments under asymmetrical pH conditions. (A) Scheme of double solution exchange protocol used for measurements under asymmetrical pH conditions. Circle represents liposome with reconstituted DgoT adsorbed on the sensor surface, at the bottom is the sequence of solution flow. (B, C) Transient currents measured in double exchange protocol experiments. Currents recorded at symmetrical pH conditions (pH 8) are shown for reference. Curves from each plot were recorded on the same sensor and therefore are directly comparable. The figure was partly adapted from [A].

became lower, while fast pre-steady state (PSS) current component dominated at pH 4.5-6. Thus, our recordings can reveal two distinct events: fast PSS reaction triggered by

52

galactonate binding (conformational change) and the slow component associated with the transport activity of the protein.

To illustrate that acidic and alkaline pH have distinct effects on symport, we tested how addition of pH gradient affects transport activity. Double solution exchange protocol was employed to generate different pH inside and outside of the vesicle (Figure 3.20A). First, sensor was rinsed and incubated with resting solution (R) without galactonate to equilibrate intraliposomal pH to the desired value. Then, single solution exchange protocol (NA-A-NA solutions) was applied, where A and NA solutions had identical pH, but different from the pH of R solution, and only A solution contained galactonate. In such configuration, during the first flow of NA solution pH gradient is established, and we can investigate its influence on galactonate transport.

Compared to symmetrical pH 8 conditions, pH 9.5 outside of the liposomes enhanced transport activity. Reversed pH gradient (i.e. pH 9.5 inside and 8 outside of the liposomes), on the other hand, significantly reduced transport (Figure 3.20B). Thus, lack of protons outside the liposomes leads to impaired proton binding and consequently lowers transport activity. Acidic pH 6 does not influence currents when applied outside the liposomes, but reduces transport when provided in intraliposomal solution (Figure 3.20C). In this case, proton release inside the liposomes becomes more challenging due to increased proton concentration inside and impairs transport. Therefore, in experiments under symmetrical pH conditions (Figure 3.19A, B), we observe the result of these two different mechanisms. Analysis of the peak currents reveals apparent pKa values for proton binding and release (Figure 3.19D).

3.2.5 Determination of symport stoichiometry

"The transport currents generated by DgoT were positive (Figure 3.19), indicating that at least two protons are transported with each galactonate molecule, as previously reported [43]. We used a reversal potential assay to determine the transport stoichiometry [73]." [A] Since galactonate transport is coupled, chemical gradient of either substrate or protons can be used as driving force. For that, the liposomes must be preloaded with buffer
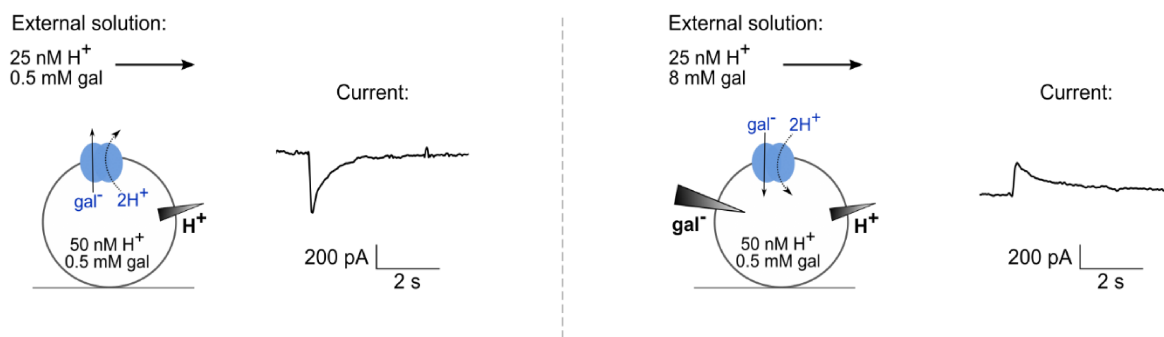
Figure 3.21. Scheme of reversal assay experiment. Different combinations of substrate and proton gradient can be used to drive coupled transport outside (*left*) or inside (*right*) the liposomes.

with known galactonate concentration, and then transport currents can be recorded using different external solutions. If only proton gradient is applied, it will solely define the direction of transport (Figure 3.21 *left*). However, if we additionally apply large enough oppositely directed substrate gradient, transport current will reverse its sign (Figure 3.21 *right*). The combination of substrate and proton gradient that results in zero transported charge depends on coupling stoichiometry, as described in Methods (eq. (2.8) and (2.10)).

In our experiment, we preloaded liposomes with buffer with pH 7.3 and 0.5 mM of galactonate (internal solution). All external solutions had pH 7.6, producing the same
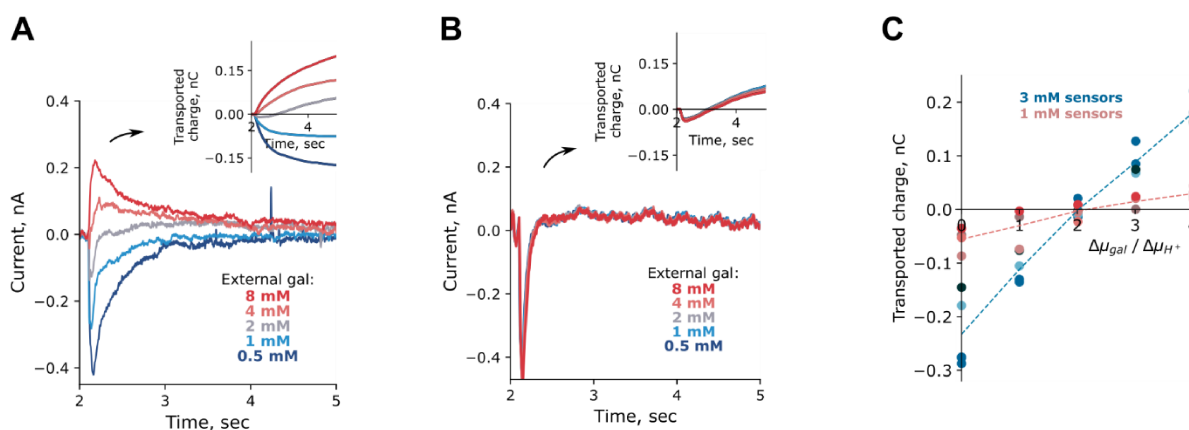
Figure 3.22. Transport stoichiometry determined with SSME. (A) Representative transport current traces used for the determination of the transport stoichiometry recorded with liposomes with DgoT using 3 mm sensors. The inset shows time dependence of transported charge obtained by integration of current traces. (B) Same as (A), but for empty liposomes. (C) Plots of transported charge *versus* ratio of the galactonate and proton chemical potentials. Values of transported charge are corrected to negative control. The figure was adapted from [A].

54

outwardly directed proton gradient. Galactonate concentration in external solutions varied between 0.5 and 8 mM, so that measured transport currents were inwardly or outwardly directed (Figure 3.22A). Generally, it is not recommended to use pH jumps in SSME experiments since they tend to induce large solution exchange artifacts [71]. However, for this assay, pH of NA and A solutions must be different to create proton concentration gradient. To address unavoidable issue of solution exchange artifact, we (i) kept pH of internal and external solutions close to each other to reduce the amplitude of artifacts and (ii) measured artifact currents with empty liposomes using the same solutions (negative control, Figure 3.22B) and corrected transported charges to the measured value. Null transport occurred at the gradient ratios corresponding to a stoichiometry of 2 H^+:1 galactonate (Figure 3.22C). Measurements using 3 mm and 1 mm sensors resulted in same x-intercept of transported charge plot.

3.2.6 Neutralization of putative proton acceptors abolishes galactonate transport

"In SSME experiments, the application of galactonate elicited a positive fast pre-steady state, but not transport currents, for D46N and E133Q DgoT, indicating that these mutant transporters can bind galactonate but cannot complete the transport cycle. Galactonate-
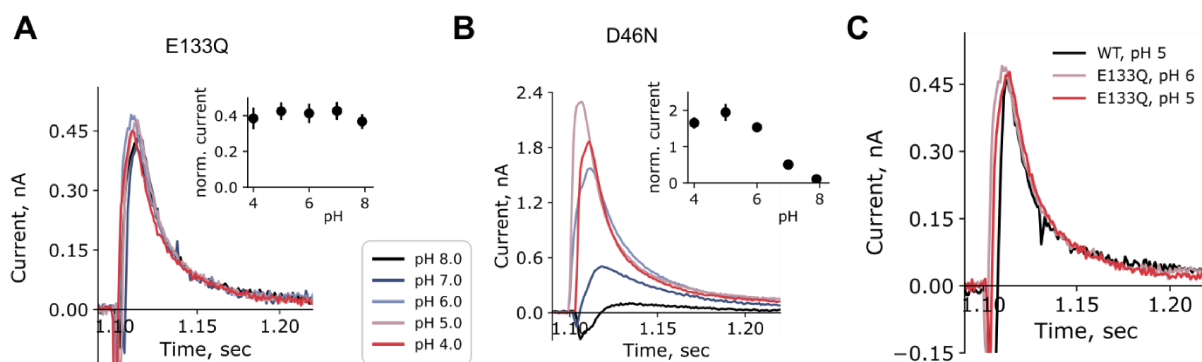
Figure 3.23. pH dependency of SSME currents for D46N and E133Q mutants. (A) SSME currents recorded with E133Q DgoT initiated by galactonate concentration jump (10 mM) at symmetrical pH conditions. (B) Same as (A), but for D46N DgoT. (C) Comparison of representative pre-steady state current recorded with WT and E133Q DgoT at pH 5 or pH 6. The figure was adapted from [A].

induced pre-steady state currents recorded with the E133Q mutant were pH independent (Figure 3.23A), suggesting that protonation of this residue is responsible for the inhibition of WT DgoT at alkaline pH. At every pH tested, E133Q currents closely resembled WT currents under acidic pH conditions (Figure 3.23C), where transport is blocked by impaired proton release inside the vesicle. In contrast, D46N currents were pH dependent, with the largest peak currents observed under acidic pH conditions (Figure 3.23B). Currents obtained with this mutant were biphasic (Figure 3.24A) and with slower decay than for E133Q. To determine whether the recorded currents represent the pre-steady state reaction or residual transport activity, we compared the time courses of current decay using liposomes reconstituted with different LPRs. Unlike the currents recorded with WT protein, the decay times for D46N DgoT did not systematically depend on the LPR (Figure 3.24B). Therefore, the observed reaction can be attributed to a slow pre-steady state process." [A]

In WT DgoT, PSS reaction was observed under extreme acidic pH, where proton release is impaired and thus transport is inhibited. This PSS current represents part of transport cycle and can be compared with partial reactions observed with different mutant variants. Analysis of PSS currents allow characterization of individual partial reactions, since the
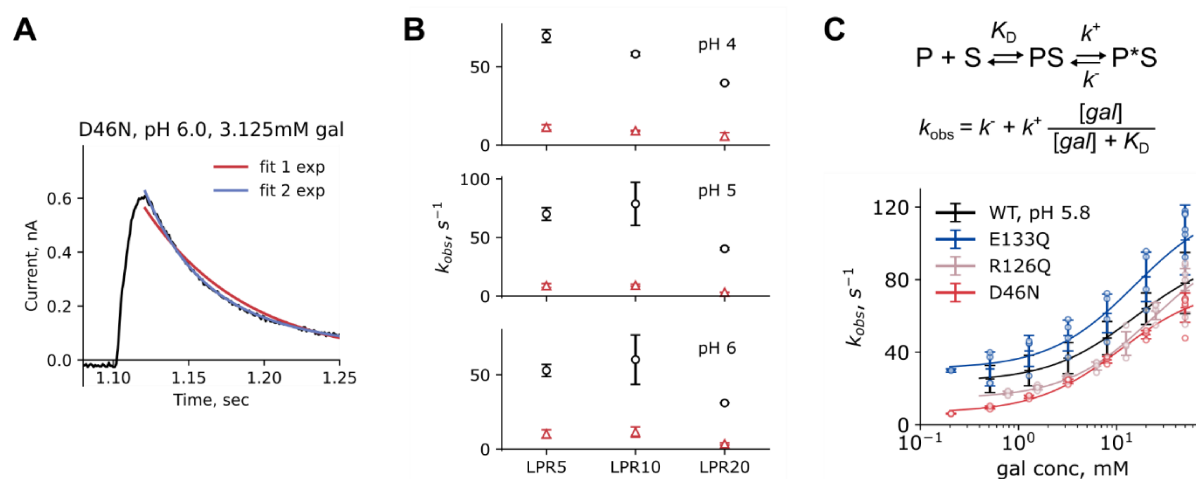
Figure 3.24. Analysis of SSME currents obtained with D46N DgoT. (A) Representative D46N DgoT current with fits to mono- (red line) or biexponential (blue line) functions. (B) Comparison of decay time constants obtained with biexponential fit of currents recorded with D46N DgoT reconstituted in liposomes at different LPR. (C) Concentration-dependent changes in k_{obs}, obtained by monoexponential fit of the current decay, for WT and mutant DgoT. Solid lines are fits to a three-states induced-fit model. Fitting equation is given above the graph. The figure was adapted from [A].

56

nature of observed reaction shapes substrate concentration dependence. If substrate binding is followed by conformational change in the protein, the observed rate constant has hyperbolic dependence on the substrate concentration (eq. (2.4)). "For D46N and E133Q DgoT, the peak current amplitudes changed with increasing galactonate concentrations in a saturating fashion, suggesting that these mutant transporters undergo conformational changes after rapid galactonate binding [88]. Similar behavior was observed for WT DgoT under acidic conditions, where only the pre-steady state reaction (no transport activity) can be seen (Figure 3.19B). Figure 3.24C shows that the observed rates (obtained as reciprocal decay time constants) depend on the galactonate concentration. Hyperbolic curve fitting provided the K_{DS} for galactonate binding, as well as the kinetic parameters characterizing the conformational changes (Table 3.2). Compared with WT, conformational changes are similar for E133Q and slower for D46N but with no difference in substrate affinity. We conclude that deprotonation of D46 and E133 is necessary to complete the transport cycle; however, partial reactions that represent conformational changes can occur in the presence of galactonate." [A]

Variant	I_{max} (nA)	k_{obs} (s^{-1}) (50 mM galactonate)	K_d (mM)	k^- (s^{-1})	k^+ (s^{-1})
WT at pH 5.8	0.7 ± 0.2	79 ± 14	15.1	23.8	70.6
D46N	2.7 ± 0.3	65 ± 8	10.9	6.6	69.9
E133Q	1.0 ± 0.3	105 ± 15	15.3	30.8	92.2
R47Q	-0.37 ± 0.05				
R126Q	-1.5 ± 0.4	74 ± 11	22	14.4	86.7

Table 3.2. Overview of kinetic parameters for transport-deficient DgoT vatiants. I_{max} values are given for recordings with 50 mM galactonate concentration jump, kinetic parameters are averaged over n = 3 different sensors (errors reflect standard deviations). For R47Q analysis of current decays was limited by time resolution of the instrument, therefore only I_{max} values are given.

3.2.7 Distinct roles of transmembrane arginine residues

As has been reported before, R47 is crucial for functional activity of DgoT and its replacement with lysine is not tolerated [42,43]. In MD simulations, R126 and R47 formed

salt bridges with D46 and E133, respectively (Figure 3.1). To better understand individual roles of these residues, we characterized experimentally DgoT variants with either arginine mutated. "For neutralizing mutations of both transmembrane arginine residues (R47Q and R126Q), galactonate application elicited fast negative currents in SSME experiments (Figure 3.25A, B). For R126Q, current amplitudes were similar at pH 5–7, but lower at more alkaline pH. Peak currents changed with increasing galactonate concentrations with a hyperbolic concentration dependence (Figure 3.24C), indicating conformational changes." [A]

For R47Q, currents were decaying faster than for other tested mutants, and analysis of decay was limited by the time resolution of the instrument. Therefore, we used peak values of measured PSS currents to describe observed electrogenic reaction. "R47Q DgoT exhibited faster currents with amplitudes increasing linearly with increasing substrate concentration (Figure 3.25C). Therefore, the observed pre-steady state reaction for R47Q differs from those observed for other transport-deficient mutants and likely represents electrogenic substrate binding rather than conformational changes [88]." [A] In this case, simple two-states model is sufficient to describe the process (Figure 3.25C).
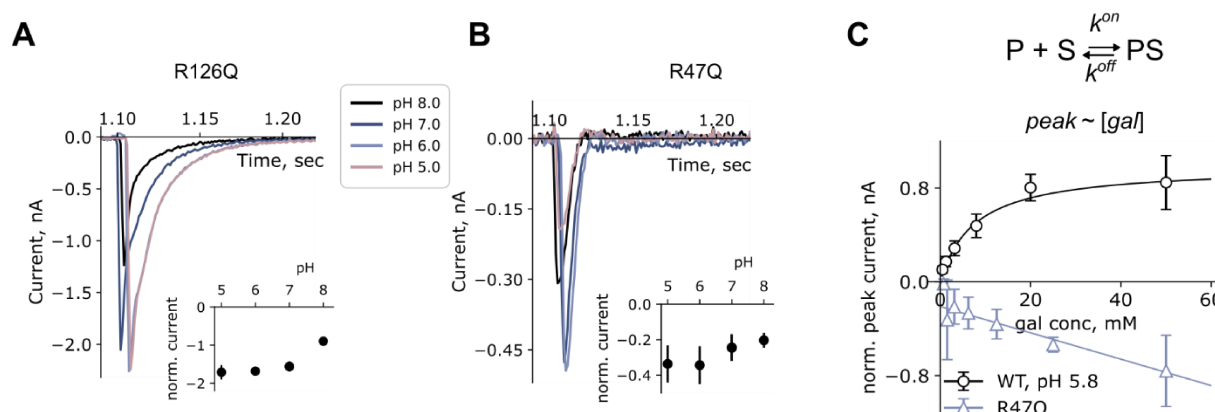
Figure 3.25. Analysis of SSME currents for arginine mutants. (A) SSME currents recorded with R126Q DgoT initiated by galactonate concentration jump (10 mM) at symmetrical pH conditions. (B) Same as (A), but for R47Q DgoT. (C) Substrate concentration dependence of the peak current values for WT and R47Q DgoT. R47Q concentration dependences were fitted to a two-state binding model given above the graph. Values of peak current values for WT DgoT, fitted to three-states model from (A) are given for comparison. The figure was adapted from [A].

3.2.8 Characterization of C43 mutations

In addition to aforementioned charged residues, mutation of C43 is able to modify functional properties of DgoT (Figure 3.26A). Replacement of cysteine with alanine abolished galactonate transport, but left substrate-induced PSS current intact (Figure 3.26B). The shape of the current resembles the ones recorded with E133Q mutant (Figure 3.23A). Mutation of C43 to serine leaved transport activity unaffected (Figure 3.26C), pointing that role of cysteine residue in this position can be fulfilled by polar but not by hydrophobic side chain.

DgoT has in total five cysteine residues, which are located far from one another and therefore cannot form disulfide bridges with each other (Figure 3.26A). Another known role for cysteine residues facilitating proton transfer [89,90]. In crystal structure, C43 is located close to E133 and therefore we hypothesize that cysteine can be a part of proton pathway during its dissociation from E133.

Interestingly, VGLUTs have serine residue at the respective position and the glutamate in a position corresponding to E133 in DgoT is also conserved in VGLUTs (Figure 1.5). Thus, the pair C43-E133 in DgoT might be functionally similar to S84-E191 in VGLUT2.
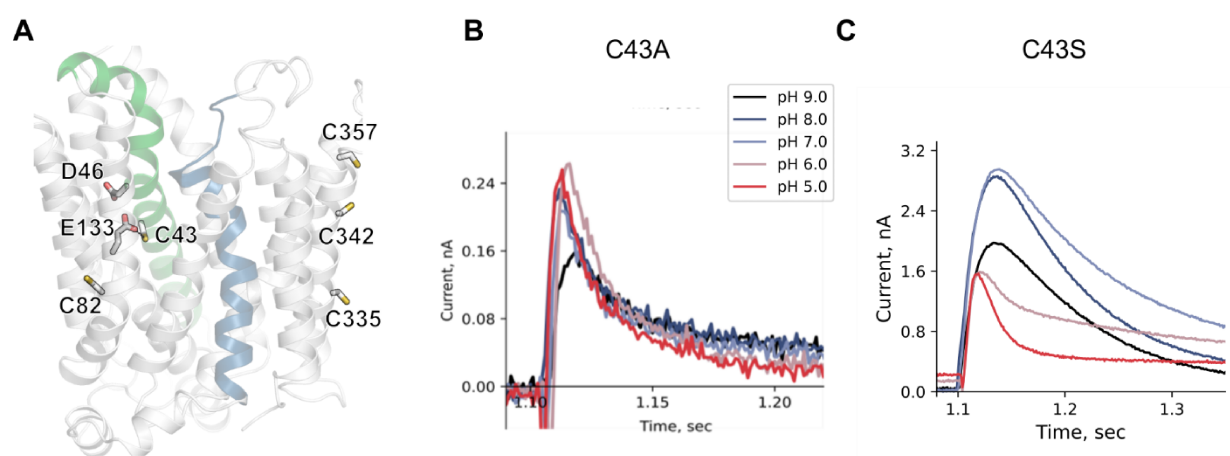
Figure 3.26. Effect of the mutations of C43. (A) Outward-facing structure with all cysteine residues of DgoT shown as sticks. TM1 is colored green and TM7 is colored blue, D46 and E133 are shown. (B) SSME currents with C43A DgoT recorded using 1 mm sensors. (C) SSME currents with C43S DgoT recorded using 3 mm sensors.

3.2.9 Substrate effect on thermal stability

Nano-differential scanning fluorimetry (nanoDSF) was used to analyze thermal stability of DgoT and effect of galactonate on it. The method is based on tryptophan fluorescence measurement. Tryptophan residues in protein absorb light at 280 nm and have variable emission spectra depending on their local environment. Emission spectrum peak shifts from 330 nm to 350 nm as protein unfolds and tryptophans become more exposed to aqueous solution. As sample is slowly heated, emission at 330 nm and 350 nm (F_{330} and F_{350}, respectively) is monitored, and their ratio indicates fraction of unfolded protein.

At 20 °C protein in the sample was folded, which resulted in lower F_{350}/F_{330} values, and as the sample was slowly heated, the ratio F_{350}/F_{330} increased (Figure 3.27A). The first derivative of the signal was used to identify inflection point more easily, as it corresponds to the position of peak in the first derivative plot.
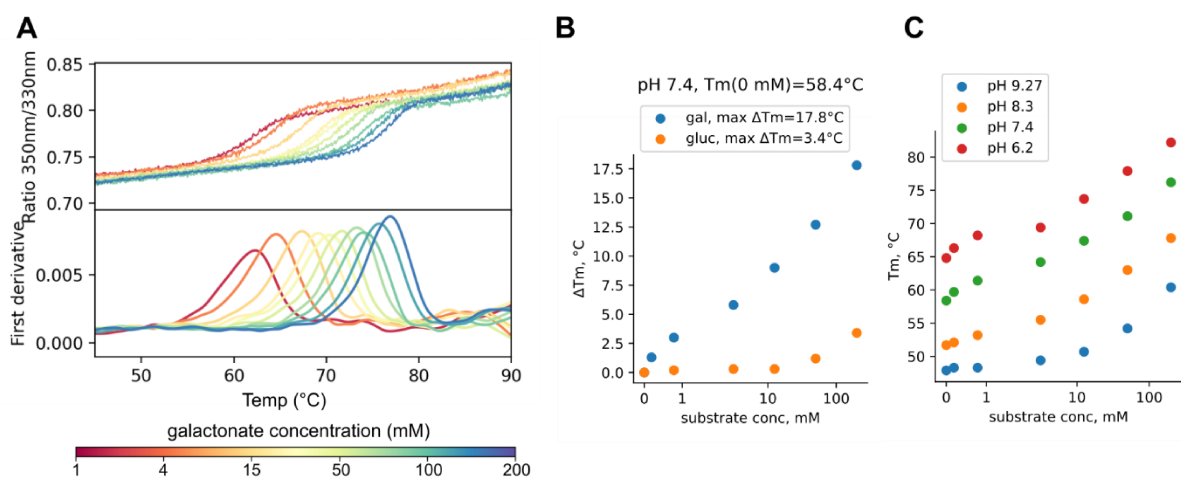
Figure 3.27. NanoDSF with WT DgoT. (A) *Top*: shifts in melting curves at different galactonate concentrations. *Bottom*: first derivative of F350/F330 ratio used to identify inflection point. (B) Shifts in T_m at different galactonate (gal) or gluconate (gluc) concentration. (C) T_m at different galactonate concentrations measured in solutions with various pH.

Addition of galactonate to purified DgoT substantially increased melting temperature of the protein (Figure 3.27A, B). Addition of gluconate up to 50 mM had no effect on thermal stability, and higher concentrations slightly increased melting temperatures (Figure 3.27B), confirming that. Thus, nanoDSF experiments demonstrate evident substrate selectivity of DgoT, in agreement with other experimental results.

pH had noticeable effect on protein stability, resulting in highest melting temperatures at pH 6.2 and lowest at pH 9.27 (Figure 3.27C). Galactonate stabilized protein at all tested pH conditions. Thermal stability in these experiments can be interpreted as conformational flexibility of the protein. At more acidic pH 6.2 WT DgoT is more thermostable, and thus, restricted flexibility reduces transport activity at these conditions (Figure 3.19B). At more alkaline pH (7.4 and 8.2) transport activity is at its maximum (Figure 3.19A), and it is associated with less thermally stable, and thus, more flexible protein.

We compared thermal stability of DgoT mutants and checked if galactonate influences it. For each protein, melting temperatures of purified protein alone and with addition of 12.5 mM of either galactonate or gluconate were measured (Figure 3.28A). At these conditions, in WT DgoT galactonate induced shift in T_m of 9 °C, while presence of gluconate did not have noticeable effect on stability. Without substrate, R126Q DgoT had lower melting temperature, E133Q and C43A variants were more stable, and the rest (D46N, R47Q and C43S) had T_m similar to the WT protein. Observed changes in protein stability can be interpreted as modified flexibility and compared with SSME results: more thermostable E133Q and C43A variants demonstrated PSS currents of lower amplitude (Figure 3.23A and Figure 3.26C) than more flexible D46N and R126 (Figure 3.23B and Figure 3.25A).
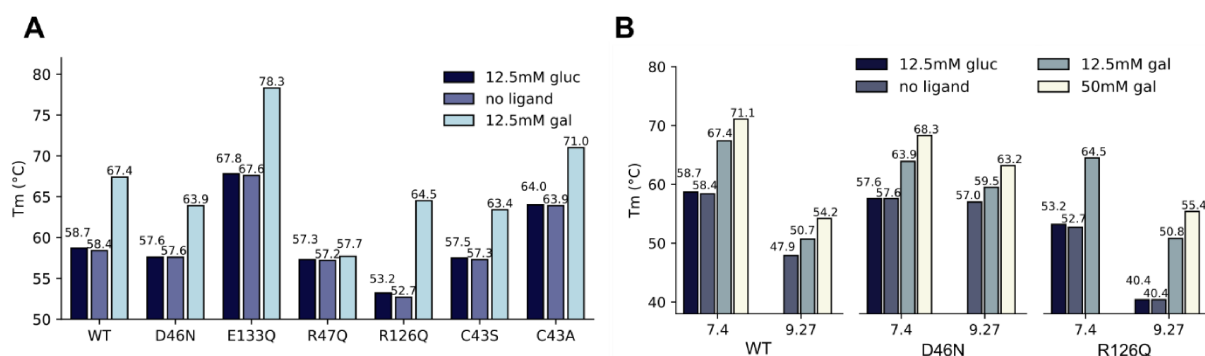
Figure 3.28. Thermal stability of DgoT mutant variants. (A) Overview of melting temperatures of different DgoT variants without substrate or in presence of galactonate (gal) or gluconate (gluc). All experiments were performed at pH 7.4. (B) Comparison of melting temperatures for WT, D46N and R126Q DgoT at neutral (7.4) and alkaline (9.27) pH.

None of the tested proteins were stabilized by gluconate, and R47Q DgoT was the only variant that did not demonstrate significantly higher T_m in presence of galactonate (Figure

3.28A). Interestingly, it is also the only mutant that did not show substrate-induced conformational changes in SSME experiments (Figure 3.25C). Thus, nanoDSF captured failure of galactonate to stabilize R47Q DgoT, in agreement with SSME results and MD simulations (Figure 3.10).

For D46N and R126Q variants, PSS currents in SSME experiments had lower amplitude at higher pH (Figure 3.23B and Figure 3.25A), therefore we additionally measured their thermostability at pH 9.27 (Figure 3.28B). WT protein had lowered T_m at alkaline pH and responded to galactonate addition with increased stability. R126Q mutant followed the same trend, while being less thermostable than WT at all tested conditions. D46N DgoT, on the other hand, was not destabilized by alkaline pH, but shifts in T_m in presence of galactonate were lowered. Thus, we can conclude that D46N DgoT can successfully bind substrate even at alkaline pH. However, conformational changes at these conditions are not electrogenic anymore, therefore PSS currents observed in SSME experiments are reduced.

4 Discussion

Members of the major facilitator superfamily are found in various cell types and fulfill diverse biological roles. To efficiently accumulate the substrates inside the cell or cell compartment, transporters often rely on coupling to another ion that is simultaneously moved down its concentration gradient. This ability distinguishes secondary active transporters from uniporters that are only capable of moving the substrate along its concentration gradient. Intriguingly, proteins seem to be able to adapt to various environmental requirements to optimize their coupling mechanism, resulting in large variability in substrate specificity and coupling stoichiometry. Thus, detailed analysis of coupling mechanism of different family members is essential for understanding of determinants of coupled transport.

Here, we investigated transport mechanism of DgoT, a bacterial D-galactonate/H$^+$ symporter that serves as a simplified model of SLC17 family transporters. We combined computational and experimental techniques to build complete transport cycle model and describe individual roles of key residues.

4.1 Galactonate transport is coupled to protons

Using SSM-based electrophysiology, we determined that DgoT transports each galactonate molecule together with 2 H$^+$ (Figure 3.22). Such stoichiometry results in positive net charge transported, despite galactonate being negatively charged at neutral pH (pK of carboxyl group is estimated to be 3.39). To ensure strict coupling, protein must be able to bind both substrate and two protons simultaneously in some intermediate states. Thus, D46 and E133 play roles of two proton binding sites that are required for functional activity of DgoT.

Notably, the two acidic residues are located close to each other in protein structure and appear to participate in transport together rather than independently. Although some parts of the transport cycle are still accessible in D46N and E133Q mutants (i.e. substrate-induced conformational changes, as demonstrated in Figure 3.24C), neither of them can

work as 1:1 H$^+$:galactonate symporter (Figure 3.16), which might be expected if protonation and deprotonation of each of the two residues was independent of another. Instead, this result suggests some cooperativity between the two and simulations provide few atomistic insights into it. One constraint is observed during substrate binding in outward-facing DgoT, where protonation of only D46 can promote conformations with open extracellular gate in *apo* simulations, but fails to consistently close the gate in substrate-bound simulations (Figure 3.1D, E). Protonation of only E133 raises the opposite problem: while extracellular gate can close in substrate-bound simulations, it opens rarely in *apo* DgoT, which might present a challenge for galactonate binding. Another restriction shows up during substrate release in inward-facing DgoT. There, deprotonation of D46 is required for galactonate release, however, most plausible pathway for the proton involves carboxyl group of E133 (Figure 3.5). Since E133 itself is protonated and cannot relay proton from D46, deprotonation of the aspartate becomes a complex process that involves consecutive proton release from E133 and from D46 via deprotonated E133.

Thus, we conclude that not only two titratable residues are required for maintaining transport stoichiometry, but also their placement in the transporter is critical.

4.2 Substrate binding is coupled to conformational changes

Galactonate transport is highly selective in DgoT [43]. It has been shown that all residues surrounding the substrate binding pocket (Y44, R47, Y79, Q164, Q264, S370 and N393) are important for active transport and mutation of even one of those reduces transport significantly [82]. In our experiments, DgoT demonstrated clear preference of galactonate over its epimer gluconate, with only the right substrate being transported (Figure 3.16). Notably, high selectivity comes together with low binding affinity. SSME experiments in non-transport conditions estimated K$_D$ of galactonate about 15 mM. The two seemingly controversial aspects can be explained by the models used to describe enzyme catalysis, where weak substrate binding to the enzyme becomes tight in the transition state [91]. Applying similar reasoning to transporters, high substrate specificity together with low binding affinity can be achieved if only correct substrate induces conformational changes necessary for the transport. Such conformational change would

lead to transition from open to occluded conformation, in which substrate binding is tighter and binding site is inaccessible from solution. Moreover, low binding affinity in open conformation can be advantageous for efficient transport, since substrate must be released at a certain point of the transport cycle, and tight binding would slow down this step.

In DgoT, the substrate plays crucial role in inducing conformational changes. In the outward-facing conformation, the extracellular gate formed by gating helices TM1 and TM7 closes when galactonate is bound, while similar molecule gluconate placed in the binding site fails to induce this conformational change (Figure 3.1). "Our combined experimental and computational results describe substrate binding as a multiphasic process initiated by galactonate recognition and accommodation in the binding site. Subsequent direct interaction between the carboxyl group of galactonate and R47 induces closure of the extracellular gate (Figure 3.10, Figure 3.11). Both reactions are electrogenic; R47Q abolishes the second step without affecting the initial substrate binding. The fast negative component corresponding to electrogenic galactonate binding also occurs with other mutants, along with an additional slower component of positive (E133Q) or negative (R126Q) amplitude (Figure 3.23A, Figure 3.25A)." [A] Thus, the organization of the binding site in DgoT is adjusted to ensure substrate selectivity as substrate must interact with several residues from different helices simultaneously to promote structural rearrangements.

Importance of substrate for conformational changes in DgoT is evident from energy landscapes of major conformational changes described with Markov state modeling. Although empty transporter can switch between inward- and outward-open conformations, energy barrier associated with transition through occluded state is lower if galactonate is bound (Figure 3.7). Moreover, presence of galactonate switches the direction of conformational changes, promoting formation of inward-facing state and thus moving further along the transport cycle.

4.3 Key amino acids play different roles in transport activity

Site-directed mutagenesis allowed us to better understand the contribution of individual residues to transport mechanism. In SSME, all investigated transport-deficient

mutants generated pre-steady state currents, indicating that some parts of transport cycle are still accessible in them. Neutralization of D46 or E133 abolishes galactonate transport due to inability of protein to maintain coupling stoichiometry (Figure 3.16). However, galactonate is still able to induce conformational changes in both mutants (Figure 3.23). According to MD simulations data, in outward-facing conformation DgoT mostly assumes occluded state in presence of substrate, regardless of D46 and E133 protonation states (Figure 3.1). Thus, parts of the transport cycle that involve galactonate are intact in D46N and E133Q mutants, while reorientation of empty carrier is not possible in them. Notably, substrate-induced conformational changes are pH-independent in E133Q mutant and have larger amplitude at acidic pH in D46N mutant. This observation indicates that PSS currents measured with these variants represent different processes, although we cannot unequivocally attribute a specific conformational change to it. In addition, such conclusion agrees with the literature, where galactonate exchange is demonstrated for E133Q, but not D46N DgoT [42]. Protonation of E133 is needed for substrate binding (Figure 3.1) and seemingly does not interfere with substrate release (Figure 3.4), therefore, E133Q mutant can readily and quickly bind and release galactonate to either side of the membrane. Deprotonation of D46, on the other hand, is needed for substrate release (Figure 3.4) and cannot be achieved in D46N mutant, making it not capable of galactonate exchange.

R47 is the residue that is crucial for coupling galactonate transport with protons in DgoT. It interacts with carboxyl group of E133 and galactonate at the same time (Figure 3.1), making it involved in substrate coordination in the binding site and the following conformational changes in the protein (Figure 3.10, Figure 3.11). Since arginine residue is conserved across SLC17 family (Figure 1.5), we hypothesize a similar role in coupling of anion transport with protons in mammalian isoforms.

Unlike R47Q, mutation of R126 does not abolish substrate-induced conformational changes (Figure 3.25). However, loss of transport activity demonstrates important role of this arginine in functional activity of DgoT. That could be, for example, regulation of D46 deprotonation in inward-facing conformation or participating directly in major conformational changes in *apo* transporter. pH dependency of measured PSS current amplitudes indicates that mutation of R126 might cause changes the pKa of D46 or E133.

Another residue involved in transport activity of the protein is C43. While mutation of this cysteine to alanine makes protein transport deficient and only capable of pH-independent pre-steady state reaction, its replacement with serine fully restores galactonate transport (Figure 3.26). We hypothesize that side chain of the cysteine mediate proton release from E133, since their side chains are positioned in vicinity of each other. Presence of serine residue in a corresponding position in VGLUTs (Figure 1.5) suggests that its role also might be related to proton transfer. Interestingly, in LacY there is no cysteine or serine residue in similar position. Instead, TM5 has who cysteine residue, C148 and C154, that appear to be important for structural integrity of the protein, but do not participate directly in active transport [92–94].

4.4 Proton transfer in MD simulations

In classical MD simulations, protonation state of titratable residues is set before the start of production run and therefore proton transfer cannot be simulated [49]. The most direct way to overcome this challenge is to use quantum mechanics/molecular mechanics (QM/MM) simulations [95]. This approach, however, comes with substantially increased computational cost of simulations. In QM/MM simulations, part of the system is treated on quantum mechanical level, while the rest is simulated classically. Due to large computational demands, system can only be simulated for shorter time, and the reliability of the results depends greatly on the quality of the initial conditions used. Therefore, a deep understanding of molecular mechanisms of the investigated system must be obtained elsewhere prior to QM/MM calculations.

Alternatively, constant pH simulations can be used to obtain pK values of titratable residues of the protein [96]. With this method, simulation time is comparable with that of classical MD. However, membrane proteins still present a challenge for the algorithm since force fields parameters were optimized for soluble proteins where titratable residues are usually well solvated. In addition, careful analysis of pK calculations is required to correctly reconstruct order of events.

To study proton coupling in DgoT, we focused our efforts on classical MD simulations. Although the method has its limitations when it comes to investigation of proton transfer

events, we can still rationalize our observations from simulations with fixed protonation states of key residues. Thus, we were able to identify coupling between protonation of DgoT and conformational changes in both outward-facing and inward-facing states (Figure 3.1, Figure 3.3). In both systems, however, additional considerations are required to understand causal relationship between them. In outward-facing state, protonation of D46 and E133 is more likely to take place via conformational selection rather than induced fit mechanism. When the extracellular gate is in closed conformation, water molecules from the bulk are separated from the active center of the protein, providing no viable path for protons (Figure 3.2A). In contrast, in the conformation with open gate active center is connected to the extracellular solution and potential pathway for protons can be found (Figure 3.2B). When both D46 and E133 are deprotonated, extracellular gate is flexible and can assume both open and closed conformations (Figure 3.1D), making protonation of the two via conformational selection mechanism feasible. In inward-facing DgoT at least one proton must be released prior to the substrate dissociation (Figure 3.3). The conformational change (i.e. opening of the intracellular gate) does not seem to be required for deprotonation of E133 (Figure 3.6A). However, deprotonation of D46 is more challenging since the only plausible pathway for H$^+$ includes carboxyl group of E133 (Figure 3.5). We analyzed how subtle conformational changes in our system modify electrostatics (Figure 3.5B, Figure 3.6B) and proposed a multi-step mechanism in which deprotonation of E133 is followed by proton transfer from D46 to E133. In this state, galactonate can be released and after follow-up deprotonation of E133 all substrates are transported. To further support our proton release model, in the paper we additionally employ QM/MM calculations to analyze the energetics of proton transfer in inward-facing DgoT [A].

4.5 Similarities and differences with other transporters

"*E. coli* can use galactonate as the sole carbon source [97], demonstrating that DgoT transports as effectively as alternative glucose transporters, such as lactose permease LacY [5] or xylose symporter XylE [98]. LacY and XylE transport uncharged sugar molecules with a stoichiometry of 1 proton:1 substrate, whereas DgoT co-transports two protons with the negatively charged galactonate, together resulting in a positive net transported charge.

Adjusting the transport stoichiometry permits DgoT to utilize ΔpH and $\Delta\Psi$ and provides high driving forces, thus optimizing bacteria for nutrient uptake when resources are limited.

Transport substrates and stoichiometries vary substantially among MFS transporters. LacY [5] and fucose transporter FucP [78] share a stoichiometry of 1 H+:1 substrate. In LacY, protonation of a single site, E325 (equivalent to E133 in DgoT [5]), permits lactose binding and transition to the occluded conformation. FucP is more similar to DgoT structurally and has two protonation sites, D46 and E135 (homologous to D46 and E133 in DgoT). However, only D46 is accessible from the extracellular solution, with E135 serving only as part of the proton transfer pathway. SLC17A5/sialin differs from DgoT in having broad substrate specificity: it recognizes various sialic acids via electrostatic interaction with two conserved arginines, R168 and R57. Two glutamates, E171 and E175, serve as protonation sites. It has been proposed that protonation of E171 releases R168 and permits its interaction with the substrate [28]. Subsequent H+ transfer to E175 is followed by substrate transfer to a more cytoplasmic binding site close to R57 and translocation to the inward-facing conformation. Such coupling of substrate translocation to proton transfer results in the electroneutral co-transport of one H+ and one sialic acid molecule.

The comparison of DgoT with FucP and SLC17A5/sialin reveals how small variations in the arrangement of protonation sites can adjust transport stoichiometries of H+-coupled transporters with conserved architecture and transport mechanisms. The SLC17 family encompasses organic anion uniporters [99,100], H+-glutamate exchangers [27], and a H+-sialic acid symporter [28]. Certain SLC17 transporters can also function as a Na+-PO_4^{3-} symporter [39,99]. Therefore, our work may serve as a framework to understand the mechanisms underlying the diversity of SLC17 transport mechanisms." [A]

5 Conclusion

The objective of this study was to generate a detailed description of transport cycle of DgoT, a proton-coupled symporter from *E.coli*. By combining MD simulations with various experimental methods, we could characterize determinants of galactonate transport. We reconstituted purified DgoT into proteoliposomes and with SSME demonstrated experimentally that galactonate transport is coupled with two protons. This refined previously presented in literature idea of several H+ co-transported with substrate [43]. At symmetrical pH conditions, transport activity is highest at a value of 7.5, close to internal pH of *E.coli* [101]. Since external pH of bacteria is usually lower, proton coupling makes transport of negatively charged galactonate more efficient.

MD simulations have shed light into some features of conformational changes. Substrate induces conformational changes in the protein by interacting with several charged and polar residues lining binding pocket. As a result, it brings two halves of the transporter together with gating helices (TM1 and TM7 from the extracellular side and TM4 and TM10 from the intracellular side) moving most prominently. Gating helices are responsible for the formation of occluded conformations and separation of binding site from the solution. We described inward-outward transition for *apo* as well as galactonate-bound DgoT and demonstrated that the latter system assumes occluded conformations and thus has lower energy barrier between oppositely facing states. *Apo* protein, on the other hand, is more flexible and visits both open and occluded conformations.

We described five residues in transmembrane region of DgoT that are involved in coupled transport. Residues D46 and E133 are protonation sites that are needed to transfer 2 H+ with each substrate molecule. Residue R126 participates in conformational changes by electrostatic interaction with deprotonated D46 and stabilizing open gate conformation in outward-facing DgoT Residue R47 from TM1 plays particularly important role in coupling substrate binding with conformational changes by forming a salt bridge with carboxyl group of galactonate and thus driving TM1 closer to TM7. Finally, C43 seems to be involved in proton transfer process and can be substituted with serine without any changes in galactonate transport efficiency.

Based on our results, we propose a detailed scheme of transport cycle of DgoT that takes into accounts symport stoichiometry, alternating access mechanism and coupling between substrates and conformational changes. This work can serve as a framework for better understanding coupled transport facilitated by SLC17 transporters.

6 Outlook

This work opens new questions about features of coupled transport mechanism in DgoT. Additional simulations on quantum mechanical level could provide better understanding of proton association and release processes. Modeling and analysis of major conformational changes in different protonation states might highlight important structural features of *apo* transporter isomerization. Also, similar computational approaches can be applied to other SLC17 members to investigate to what extend transport mechanism of DgoT resembles those of mammalian proteins. Further experiments could provide more support to computational findings. Examination of conformational state of the protein in experiment and real-time monitoring of conformational changes, for example, using fluorescent methods, could confirm induced fit mechanism of substrate binding. Mutagenesis of hydrophobic residues from neck regions of gating helices could highlight their role in tight coupling of galactonate transport with protons.

7 Appendix

Structure	D46 protonation	E133 protonation	substrate	# replicas	Total simulation time (µs)
WT, outward	-	-	no	8	3.5
	-	H	no	4	0.8
	H	-	no	4	0.8
	H	H	no	4	0.8
	-	H	100 mM galactonate in bulk	5	3.2
	H	-		5	2.7
	H	H		5	3.2
	-	-	Bound galactonate[-]	5	1.3
	-	H		4	1.5
	H	-		4	2.8
	H	H		4	2.0
	H	H	Bound gluconate	5	2.0
WT, inward	-	-	no	4	1.7
	-	H	no	4	0.7
	H	H	no	4	1.8
	-	-	Bound galactonate[-]	4	2.6
	-	H		10	3.6
	H	-		10	3.9
	H	H		5	2.0
	-	H	Bound galactonate[H]	4	2.0
	H	-		4	1.5
R47Q, outward	H	H	no	4	2.0
	H	H	Bound galactonate[-]	4	2.0

Table 7.1. Summary of unbiased MD simulations with WT and mutant DgoT initiated from crystal structures.

8 Bibliography

[1] Drew D, North RA, Nagarathinam K, Tanabe M. Structures and General Transport Mechanisms by the Major Facilitator Superfamily (MFS). Chem Rev 2021;121:5289–335. https://doi.org/10.1021/acs.chemrev.0c00983.

[2] Quistgaard EM, Löw C, Guettou F, Nordlund P. Understanding transport by the major facilitator superfamily (MFS): structures pave the way. Nat Rev Mol Cell Biol 2016;17:123–32. https://doi.org/10.1038/nrm.2015.25.

[3] Abramson J, Iwata S, Kaback HR. Lactose permease as a paradigm for membrane transport proteins (Review). Mol Membr Biol 2004;21:227–36. https://doi.org/10.1080/09687680410001716862.

[4] Madej MG. Function, Structure, and Evolution of the Major Facilitator Superfamily: The LacY Manifesto. Advances in Biology 2014;2014:1–20. https://doi.org/10.1155/2014/523591.

[5] Kaback HR, Guan L. It takes two to tango: The dance of the permease. Journal of General Physiology 2019;151:878–86. https://doi.org/10.1085/jgp.201912377.

[6] Abramson J, Smirnova I, Kasho V, Verner G, Kaback HR, Iwata S. Structure and Mechanism of the Lactose Permease of Escherichia coli. Science 2003;301:610–5. https://doi.org/10.1126/science.1088196.

[7] Sugihara J, Smirnova I, Kasho V, Kaback HR. Sugar recognition by CscB and LacY. Biochemistry 2011;50:11009–14. https://doi.org/10.1021/bi201592y.

[8] Guan L, Kaback HR. Binding affinity of lactose permease is not altered by the H+ electrochemical gradient. Proc Natl Acad Sci U S A 2004;101:12148–52. https://doi.org/10.1073/pnas.0404936101.

[9] Smirnova I, Kasho V, Kaback HR. Real-time conformational changes in LacY. Proc Natl Acad Sci U S A 2014;111:8440–5. https://doi.org/10.1073/pnas.1408374111.

[10] Kumar H, Finer-Moore JS, Kaback HR, Stroud RM. Structure of LacY with an α-substituted galactoside: Connecting the binding site to the protonation site. Proceedings of the National Academy of Sciences 2015;112:9004–9. https://doi.org/10.1073/pnas.1509854112.

[11] Kumar H, Finer-Moore J, Smirnova I, Kasho V, Pardon E, Steyaert J, et al. Diversity in kinetics correlated with structure in nano body-stabilized LacY. PLOS ONE 2020;15:e0232846. https://doi.org/10.1371/journal.pone.0232846.

[12] Smirnova I, Kasho V, Sugihara J, Vázquez-Ibar JL, Kaback HR. Role of protons in sugar binding to LacY. Proc Natl Acad Sci U S A 2012;109:16835–40. https://doi.org/10.1073/pnas.1214890109.

[13] Frillingos S, Sahin-Tóth M, Wu J, Kaback HR. Cys-scanning mutagenesis: a novel approach to structure—function relationships in polytopic membrane proteins. The FASEB Journal 1998;12:1281–99. https://doi.org/10.1096/fasebj.12.13.1281.

[14] Grytsyk N, Sugihara J, Kaback HR, Hellwig P. pKa of Glu325 in LacY. Proc Natl Acad Sci U S A 2017;114:1530–5. https://doi.org/10.1073/pnas.1621431114.

[15] Carrasco N, Püttner IB, Antes LM, Lee JA, Larigan JD, Lolkema JS, et al. Characterization of site-directed mutants in the lac permease of Escherichia coli. 2.

Glutamate-325 replacements. Biochemistry 1989;28:2533–9. https://doi.org/10.1021/bi00432a028.

[16]   Garcia-Celma JJ, Ploch J, Smirnova I, Kaback HR, Fendler K. Delineating electrogenic reactions during lactose/H+ symport. Biochemistry 2010;49:6115–21. https://doi.org/10.1021/bi100492p.

[17]   Sahin-Toth M, Kaback HR. Arg-302 facilitates deprotonation of Glu-325 in the transport mechanism of the lactose permease from Escherichiacoli. Proc Natl Acad Sci U S A 2001;98:6068–73. https://doi.org/10.1073/pnas.111139698.

[18]   Weinglass AB, Smirnova IN, Kaback HR. Engineering Conformational Flexibility in the Lactose Permease of Escherichia coli: Use of Glycine-Scanning Mutagenesis To Rescue Mutant Glu325→Asp. Biochemistry 2001;40:769–76. https://doi.org/10.1021/bi002171m.

[19]   Grytsyk N, Santos Seiça AF, Sugihara J, Kaback HR, Hellwig P. Arg302 governs the pKa of Glu325 in LacY. Proc Natl Acad Sci U S A 2019;116:4934–9. https://doi.org/10.1073/pnas.1820744116.

[20]   Kaback HR, Smirnova I, Kasho V, Nie Y, Zhou Y. The Alternating Access Transport Mechanism in LacY. J Membr Biol 2011;239:85–93. https://doi.org/10.1007/s00232-010-9327-5.

[21]   Forrest LR, Krämer R, Ziegler C. The structural basis of secondary active transport mechanisms. Biochim Biophys Acta 2011;1807:167–88. https://doi.org/10.1016/j.bbabio.2010.10.014.

[22]   Zhang XC, Zhao Y, Heng J, Jiang D. Energy coupling mechanisms of MFS transporters. Protein Sci 2015;24:1560–79. https://doi.org/10.1002/pro.2759.

[23]   Qureshi AA, Suades A, Matsuoka R, Brock J, McComas SE, Nji E, et al. The molecular basis for sugar import in malaria parasites. Nature 2020;578:321–5. https://doi.org/10.1038/s41586-020-1963-z.

[24]   Li F, Eriksen J, Finer-Moore J, Stroud RM, Edwards RH. Diversity of function and mechanism in a family of organic anion transporters. Current Opinion in Structural Biology 2022;75:102399. https://doi.org/10.1016/j.sbi.2022.102399.

[25]   Omote H, Miyaji T, Hiasa M, Juge N, Moriyama Y. Structure, Function, and Drug Interactions of Neurotransmitter Transporters in the Postgenomic Era. Annu Rev Pharmacol Toxicol 2016;56:385–402. https://doi.org/10.1146/annurev-pharmtox-010814-124816.

[26]   Reimer RJ. SLC17: A functionally diverse family of organic anion transporters. Molecular Aspects of Medicine 2013;34:350–9. https://doi.org/10.1016/j.mam.2012.05.004.

[27]   Kolen B, Borghans B, Kortzak D, Lugo V, Hannack C, Guzman RE, et al. Vesicular glutamate transporters are H+-anion exchangers that operate at variable stoichiometry. Nat Commun 2023;14:2723. https://doi.org/10.1038/s41467-023-38340-9.

[28]   Hu W, Chi C, Song K, Zheng H. The molecular mechanism of sialic acid transport mediated by Sialin. Science Advances 2023;9:eade8346. https://doi.org/10.1126/sciadv.ade8346.

[29]   Morin P, Sagné C, Gasnier B. Functional characterization of wild-type and mutant human sialin. EMBO J 2004;23:4560–70. https://doi.org/10.1038/sj.emboj.7600464.

[30]   Pietrancosta N, Djibo M, Daumas S, El Mestikawy S, Erickson JD. Molecular, Structural, Functional, and Pharmacological Sites for Vesicular Glutamate Transporter

Regulation. Mol Neurobiol 2020;57:3118–42. https://doi.org/10.1007/s12035-020-01912-7.

[31]   Fremeau RT, Troyer MD, Pahner I, Nygaard GO, Tran CH, Reimer RJ, et al. The expression of vesicular glutamate transporters defines two classes of excitatory synapse. Neuron 2001;31:247–60. https://doi.org/10.1016/s0896-6273(01)00344-0.

[32]   Schäfer MK-H, Varoqui H, Defamie N, Weihe E, Erickson JD. Molecular Cloning and Functional Identification of Mouse Vesicular Glutamate Transporter 3 and Its Expression in Subsets of Novel Excitatory Neurons*. Journal of Biological Chemistry 2002;277:50734–48. https://doi.org/10.1074/jbc.M206738200.

[33]   Kashani A, Betancur C, Giros B, Hirsch E, Mestikawy SE. Altered expression of vesicular glutamate transporters VGLUT1 and VGLUT2 in Parkinson disease. Neurobiology   of   Aging   2007;28:568–78. https://doi.org/10.1016/j.neurobiolaging.2006.02.010.

[34]   Kashani A, Lepicard È, Poirel O, Videau C, David JP, Fallet-Bianco C, et al. Loss of VGLUT1 and VGLUT2 in the prefrontal cortex is correlated with cognitive decline in Alzheimer   disease.   Neurobiology   of   Aging   2008;29:1619–30. https://doi.org/10.1016/j.neurobiolaging.2007.04.010.

[35]   Varoqui H, Schäfer MKH, Zhu H, Weihe E, Erickson JD. Identification of the differentiation-associated Na+/PI transporter as a novel vesicular glutamate transporter expressed in a distinct set of glutamatergic synapses. J Neurosci 2002;22:142–55. https://doi.org/10.1523/JNEUROSCI.22-01-00142.2002.

[36]   Juge N, Yoshida Y, Yatsushiro S, Omote H, Moriyama Y. Vesicular Glutamate Transporter Contains Two Independent Transport Machineries *. Journal of Biological Chemistry 2006;281:39499–506. https://doi.org/10.1074/jbc.M607670200.

[37]   Preobraschenski J, Zander J-F, Suzuki T, Ahnert-Hilger G, Jahn R. Vesicular Glutamate Transporters Use Flexible Anion and Cation Binding Sites for Efficient Accumulation of Neurotransmitter.   Neuron   2014;84:1287–301. https://doi.org/10.1016/j.neuron.2014.11.008.

[38]   Chang R, Eriksen J, Edwards RH. The dual role of chloride in synaptic vesicle glutamate transport. eLife 2018;7:e34896. https://doi.org/10.7554/eLife.34896.

[39]   Preobraschenski J, Cheret C, Ganzella M, Zander JF, Richter K, Schenck S, et al. Dual and Direction-Selective Mechanisms of Phosphate Transport by the Vesicular Glutamate   Transporter.   Cell   Rep   2018;23:535–45. https://doi.org/10.1016/j.celrep.2018.03.055.

[40]   Takamori S. Vesicular glutamate transporters as anion channels? Pflugers Arch 2016;468:513–8. https://doi.org/10.1007/s00424-015-1760-y.

[41]   Aula N, Salomäki P, Timonen R, Verheijen F, Mancini G, Månsson J-E, et al. The Spectrum of SLC17A5-Gene Mutations Resulting in Free Sialic Acid–Storage Diseases Indicates Some Genotype-Phenotype Correlation. The American Journal of Human Genetics 2000;67:832–40. https://doi.org/10.1086/303077.

[42]   Batarni S, Nayak N, Chang A, Li F, Hareendranath S, Zhou L, et al. Substrate recognition and proton coupling by a bacterial member of solute carrier family 17. J Biol Chem 2023;299:104646. https://doi.org/10.1016/j.jbc.2023.104646.

[43]   Leano JB, Batarni S, Eriksen J, Juge N, Pak JE, Kimura-Someya T, et al. Structures suggest a mechanism for energy coupling by a family of organic anion transporters. PLoS Biol 2019;17:e3000260. https://doi.org/10.1371/journal.pbio.3000260.

[44]   Li F, Eriksen J, Finer-Moore J, Chang R, Nguyen P, Bowen A, et al. Ion transport and regulation in a synaptic vesicle glutamate transporter. Science 2020;368:893–7. https://doi.org/10.1126/science.aba9202.

[45]   Klingenberg M. Transport viewed as a catalytic process. Biochimie 2007;89:1042–8. https://doi.org/10.1016/j.biochi.2007.02.010.

[46]   Kaback HR. A chemiosmotic mechanism of symport. Proc Natl Acad Sci U S A 2015;112:1259–64. https://doi.org/10.1073/pnas.1419325112.

[47]   Vogt AD, Di Cera E. Conformational selection or induced fit? A critical appraisal of the kinetic mechanism. Biochemistry 2012;51:5894–902. https://doi.org/10.1021/bi3006913.

[48]   Hammes GG, Chang Y-C, Oas TG. Conformational selection or induced fit: A flux description of reaction mechanism. Proceedings of the National Academy of Sciences 2009;106:13737–41. https://doi.org/10.1073/pnas.0907195106.

[49]   Hollingsworth SA, Dror RO. Molecular dynamics simulation for all. Neuron 2018;99:1129–43. https://doi.org/10.1016/j.neuron.2018.08.011.

[50]   Amadei A, Linssen AB, Berendsen HJ. Essential dynamics of proteins. Proteins 1993;17:412–25. https://doi.org/10.1002/prot.340170408.

[51]   Amadei A, Linssen AB, de Groot BL, van Aalten DM, Berendsen HJ. An efficient method for sampling the essential subspace of proteins. J Biomol Struct Dyn 1996;13:615–25. https://doi.org/10.1080/07391102.1996.10508874.

[52]   Waterhouse A, Bertoni M, Bienert S, Studer G, Tauriello G, Gumienny R, et al. SWISS-MODEL: homology modelling of protein structures and complexes. Nucleic Acids Res 2018;46:W296–303. https://doi.org/10.1093/nar/gky427.

[53]   Schrödinger, LLC. The PyMOL Molecular Graphics System, Version 1.8 2015.

[54]   Lomize MA, Pogozheva ID, Joo H, Mosberg HI, Lomize AL. OPM database and PPM web server: resources for positioning of proteins in membranes. Nucleic Acids Res 2012;40:D370-376. https://doi.org/10.1093/nar/gkr703.

[55]   Wolf MG, Hoefling M, Aponte-Santamaría C, Grubmüller H, Groenhof G. g_membed: Efficient insertion of a membrane protein into an equilibrated lipid bilayer with minimal perturbation. J Comput Chem 2010;31:2169–74. https://doi.org/10.1002/jcc.21507.

[56]   Abraham MJ, Murtola T, Schulz R, Páll S, Smith JC, Hess B, et al. GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. SoftwareX 2015;1–2:19–25. https://doi.org/10.1016/j.softx.2015.06.001.

[57]   Huang J, Rauscher S, Nawrocki G, Ran T, Feig M, de Groot BL, et al. CHARMM36m: an improved force field for folded and intrinsically disordered proteins. Nat Methods 2017;14:71–3. https://doi.org/10.1038/nmeth.4067.

[58]   Zoete V, Cuendet MA, Grosdidier A, Michielin O. SwissParam: a fast force field generation tool for small organic molecules. J Comput Chem 2011;32:2359–68. https://doi.org/10.1002/jcc.21816.

[59]   Essmann U, Perera L, Berkowitz ML, Darden T, Lee H, Pedersen LG. A smooth particle mesh Ewald method. The Journal of Chemical Physics 1995;103:8577–93. https://doi.org/10.1063/1.470117.

[60]   Bussi G, Donadio D, Parrinello M. Canonical sampling through velocity rescaling. J Chem Phys 2007;126:014101. https://doi.org/10.1063/1.2408420.

[61] Berendsen HJC, Postma JPM, van Gunsteren WF, DiNola A, Haak JR. Molecular dynamics with coupling to an external bath. The Journal of Chemical Physics 1984;81:3684–90. https://doi.org/10.1063/1.448118.

[62] Parrinello M, Rahman A. Polymorphic transitions in single crystals: A new molecular dynamics method. J Appl Phys; (United States) 1981;52:7182–90. https://doi.org/10.1063/1.328693.

[63] Pérez-Hernández G, Paul F, Giorgino T, De Fabritiis G, Noé F. Identification of slow molecular order parameters for Markov model construction. The Journal of Chemical Physics 2013;139:015102. https://doi.org/10.1063/1.4811489.

[64] Scherer MK, Trendelkamp-Schroer B, Paul F, Pérez-Hernández G, Hoffmann M, Plattner N, et al. PyEMMA 2: A Software Package for Estimation, Validation, and Analysis of Markov Models. J Chem Theory Comput 2015;11:5525–42. https://doi.org/10.1021/acs.jctc.5b00743.

[65] Deuflhard P, Weber M. Robust Perron cluster analysis in conformation dynamics. Linear Algebra and Its Applications 2005;398:161–84. https://doi.org/10.1016/j.laa.2004.10.026.

[66] Kostritskii AY, Alleva C, Cönen S, Machtens J-P. g_elpot: A Tool for Quantifying Biomolecular Electrostatics from Molecular Dynamics Trajectories. J Chem Theory Comput 2021;17:3157–67. https://doi.org/10.1021/acs.jctc.0c01246.

[67] Kostritskii AY, Machtens J-P. Domain- and state-specific shape of the electric field tunes voltage sensing in voltage-gated sodium channels. Biophysical Journal 2023;122:1807–21. https://doi.org/10.1016/j.bpj.2023.04.013.

[68] Hedenburg OF. ON THE ESTERS, AS WELL AS THE MONOMOLECULAR β- AND γ-LACTONES, OF d-MANNONIC AND d-GLUCONIC ACIDS; ON ORTHO-BIS-d-GALACTONIC ACID, d-GALACTONIC γ-LACTONE AND ITS MONO-HYDRATE. ACS Publications 2002. https://doi.org/10.1021/ja02271a012.

[69] Liu H, Valdehuesa KNG, Nisola GM, Ramos KRM, Chung W-J. High yield production of d-xylonic acid from d-xylose using engineered Escherichia coli. Bioresource Technology 2012;115:244–8. https://doi.org/10.1016/j.biortech.2011.08.065.

[70] Bosshart PD, Kalbermatter D, Bonetti S, Fotiadis D. Mechanistic basis of L-lactate transport in the SLC16 solute carrier family. Nat Commun 2019;10:2649. https://doi.org/10.1038/s41467-019-10566-6.

[71] Bazzone A, Barthmes M, Fendler K. SSM-Based Electrophysiology for Transporter Research. Methods Enzymol 2017;594:31–83. https://doi.org/10.1016/bs.mie.2017.05.008.

[72] Smirnova IN, Kasho VN, Kaback HR. Direct sugar binding to LacY measured by resonance energy transfer. Biochemistry 2006;45:15279–87. https://doi.org/10.1021/bi061632m.

[73] Thomas NE, Feng W, Henzler-Wildman KA. A solid-supported membrane electrophysiology assay for efficient characterization of ion-coupled transport. J Biol Chem 2021;297:101220. https://doi.org/10.1016/j.jbc.2021.101220.

[74] A. Senisterra G, Patrick J. Finerty J. High throughput methods of assessing protein stability and aggregation. Molecular BioSystems 2009;5:217–23. https://doi.org/10.1039/B814377C.

[75] Pantoliano MW, Petrella EC, Kwasnoski JD, Lobanov VS, Myslik J, Graf E, et al. High-density miniaturized thermal shift assays as a general strategy for drug discovery. J Biomol Screen 2001;6:429–40. https://doi.org/10.1177/108705710100600609.

[76] Humphrey W, Dalke A, Schulten K. VMD: visual molecular dynamics. J Mol Graph 1996;14:33–8, 27–8. https://doi.org/10.1016/0263-7855(96)00018-5.

[77] Feng J, Selvam B, Shukla D. How do antiporters exchange substrates across the cell membrane? An atomic-level description of the complete exchange cycle in NarK. Structure 2021;29:922-933.e3. https://doi.org/10.1016/j.str.2021.03.014.

[78] Dang S, Sun L, Huang Y, Lu F, Liu Y, Gong H, et al. Structure of a fucose transporter in an outward-open conformation. Nature 2010;467:734–8. https://doi.org/10.1038/nature09406.

[79] Kumar H, Finer-Moore JS, Kaback HR, Stroud RM. Structure of LacY with an α-substituted galactoside: Connecting the binding site to the protonation site. Proc Natl Acad Sci USA 2015;112:9004–9. https://doi.org/10.1073/pnas.1509854112.

[80] Pedersen BP, Kumar H, Waight AB, Risenmay AJ, Roe-Zurz Z, Chau BH, et al. Crystal structure of a eukaryotic phosphate transporter. Nature 2013;496:533–6. https://doi.org/10.1038/nature12042.

[81] Iancu CV, Zamoon J, Woo SB, Aleshin A, Choe J. Crystal structure of a glucose/H+ symporter and its mechanism of action. Proceedings of the National Academy of Sciences 2013;110:17862–7. https://doi.org/10.1073/pnas.1311485110.

[82] Batarni S, Nayak N, Chang A, Li F, Hareendranath S, Zhou L, et al. Substrate recognition and proton coupling by a bacterial member of solute carrier family 17. Journal of Biological Chemistry 2023:104646. https://doi.org/10.1016/j.jbc.2023.104646.

[83] Swanson JM. Multiscale kinetic analysis of proteins. Curr Opin Struct Biol 2022;72:169–75. https://doi.org/10.1016/j.sbi.2021.11.005.

[84] Stockbridge RB, Tsai M-F. Lipid reconstitution and recording of recombinant ion channels. Methods Enzymol 2015;556:385–404. https://doi.org/10.1016/bs.mie.2014.12.028.

[85] Soskine M, Adam Y, Schuldiner S. Direct evidence for substrate-induced proton release in detergent-solubilized EmrE, a multidrug transporter. J Biol Chem 2004;279:9951–5. https://doi.org/10.1074/jbc.M312853200.

[86] Bazzone A, Zabadne AJ, Salisowski A, Madej MG, Fendler K. A Loose Relationship: Incomplete H+/Sugar Coupling in the MFS Sugar Transporter GlcP. Biophys J 2017;113:2736–49. https://doi.org/10.1016/j.bpj.2017.09.038.

[87] Deamer DW. Proton permeation of lipid bilayers. J Bioenerg Biomembr 1987;19:457–79. https://doi.org/10.1007/BF00770030.

[88] Bazzone A, Tesmer L, Kurt D, Kaback HR, Fendler K, Madej MG. Investigation of sugar binding kinetics of the E. coli sugar/H+ symporter XylE using solid-supported membrane-based electrophysiology. J Biol Chem 2021;298:101505. https://doi.org/10.1016/j.jbc.2021.101505.

[89] Huynh MHV, Meyer TJ. Proton-Coupled Electron Transfer. Chem Rev 2007;107:5004–64. https://doi.org/10.1021/cr0500030.

[90] Wiedemann C, Kumar A, Lang A, Ohlenschläger O. Cysteines and Disulfide Bonds as Structure-Forming Units: Insights From Different Domains of Life and the Potential for

Characterization by NMR. Front Chem 2020;8. https://doi.org/10.3389/fchem.2020.00280.

[91]   Arcus VL, Prentice EJ, Hobbs JK, Mulholland AJ, Van der Kamp MW, Pudney CR, et al. On the Temperature Dependence of Enzyme-Catalyzed Rates. Biochemistry 2016;55:1681–8. https://doi.org/10.1021/acs.biochem.5b01094.

[92]   Trumble WR, Viitanen PV, Sarkar HK, Poonian MS, Kaback HR. Site-directed mutagenesis of cys148 in the lac carrier protein of Escherichia coli. Biochemical and Biophysical Research Communications 1984;119:860–7. https://doi.org/10.1016/0006-291X(84)90853-2.

[93]   Properties of a mutant lactose carrier of Escherichia coli with a Cys148 → Ser148 substitution. FEBS Letters 1985;185:83–8. https://doi.org/10.1016/0014-5793(85)80745-6.

[94]   Nie Y, Sabetfard FE, Kaback HR. The Cys154 → Gly Mutation in LacY Causes Constitutive Opening of the Hydrophilic Periplasmic Pathway. Journal of Molecular Biology 2008;379:695–703. https://doi.org/10.1016/j.jmb.2008.04.015.

[95]   Senn HM, Thiel W. QM/MM methods for biomolecular systems. Angew Chem Int Ed Engl 2009;48:1198–229. https://doi.org/10.1002/anie.200802019.

[96]   Aho N, Buslaev P, Jansen A, Bauer P, Groenhof G, Hess B. Scalable Constant pH Molecular Dynamics in GROMACS. J Chem Theory Comput 2022;18:6148–60. https://doi.org/10.1021/acs.jctc.2c00516.

[97]   Deacon J, Cooper RA. D-Galactonate utilisation by enteric bacteria. The catabolic pathway in Escherichia coli. FEBS Lett 1977;77:201–5. https://doi.org/10.1016/0014-5793(77)80234-2.

[98]   Madej MG, Sun L, Yan N, Kaback HR. Functional architecture of MFS D-glucose transporters. Proc Natl Acad Sci U S A 2014;111:E719-727. https://doi.org/10.1073/pnas.1400336111.

[99]   Iharada M, Miyaji T, Fujimoto T, Hiasa M, Anzai N, Omote H, et al. Type 1 sodium-dependent phosphate transporter (SLC17A1 Protein) is a Cl(-)-dependent urate exporter. J Biol Chem 2010;285:26107–13. https://doi.org/10.1074/jbc.M110.122721.

[100]   Ishikawa T, Nishikawa H, Gao Y, Sawa Y, Shibata H, Yabuta Y, et al. The pathway via D-galacturonate/L-galactonate is significant for ascorbate biosynthesis in Euglena gracilis: identification and functional characterization of aldonolactonase. J Biol Chem 2008;283:31133–41. https://doi.org/10.1074/jbc.M803930200.

[101]   Slonczewski JL, Rosen BP, Alger JR, Macnab RM. pH homeostasis in Escherichia coli: measurement by 31P nuclear magnetic resonance of methylphosphonate and phosphate. Proc Natl Acad Sci U S A 1981;78:6271–5. https://doi.org/10.1073/pnas.78.10.6271.

9 Acknowledgements

First of all, I would like to thank my supervisor Prof. Dr. Christoph Fahlke for giving me an opportunity to work on an exciting project and freedom to shape its result. I am grateful to him for providing all the necessary resources, fruitful discussions and continued support over these years.

Next, I would like to thank the members of the Institute of Neurosciences and Medicine (INM-9) Prof. Dr. Paolo Carloni, Dr. Mercedes Alfonso-Prieto and a former member Dr. Samira Gholami for their contribution to this project as collaborators.

I would also like to thank Prof. Karl-Erich Jäger for his willingness to be a referee for this dissertation.

I thank Prof. Jan-Philipp Machtens for providing computational resources and technical support. I am grateful to the members of his group Drs. Andrei Kostritskii, Bassam Haddad, Felix Groß-Esser and Piersilvio Longo for their help with computational part of this project.

I am immensely grateful to Dr. Claudia Alleva for her invaluable contribution as my supervisor. I also thank Dr. Daniel Kortzak for his help with my first steps in the molecular dynamics simulations field.

I wish to thank Meike Berndt for her help with the experimental part of this project.

Finally, I thank all current and former members of IBI-1 for making the atmosphere in this institute truly friendly and supportive.

10 Abbreviations

APS   ammonium persulfate

BCIP   5-Bromo-4-chloro-3-indolyl phosphate

Cryo-EM   cryo-electron microscopy

CV   column volume

DDM   n-dodecyl-D-maltoside

DgoT   D-galactonate transporter from *Escherichia coli*

DSF   differential scanning fluorimetry

E.coli   *Escherichia coli*

EDTA   ethylenediaminetetraacetic acid

FucP   fucose transporter from *Escherichia coli*

HEPES   4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid

IPTG   isopropyl β-d-1-thiogalactopyranoside

LacY   lactose permease

LMNG   Lauryl maltose neopentyl glycol

LPR   lipid-to-protein ratio

MD   molecular dynamics

MFS   major facilitator superfamily

MSM   Markov state model

MWCO   molecular weight cut-off

NBT   nitro blue tetrazolium

NPG   4-nitrophenyl-α-d-galactopyranoside

PAGE   polyacrylamide gel electrophoresis

PCCA   Perron-cluster cluster analysis

PCR   polymerase chain reaction

PSS   pre-steady state

RT   room temperature

SDS   sodium dodecyl sulfate

SDS-PAGE sodium dodecyl sulfate-polyacrylamide gel electrophoresis

SEC   size-exclusion chromatography

SLC   solute carrier

SSME   solid-supported membrane electrophysiology

TDG   β-d-galactopyranosyl-1-thio-β-d-galactopyranoside

TEMED   N,N,N',N'-Tetramethyl-ethylenediamine

TICA   time-lagged independent component analysis

VGLUT   vesicular glutamate transporter

WB   western blot

WT   wild type

Δμ   electrochemical gradient

ΔpH H^+ concentration gradient

ΔΨ   membrane potential

11 Eidesstattliche Versicherung

Ich versichere an Eides Statt, dass die Dissertation von mir selbstständig und ohne unzulässige fremde Hilfe unter Beachtung der "Grundsatze zur Sicherung guter wissenschaftlicher Praxis an der Heinrich-Heine-Universität Düsseldorf" erstellt worden ist.

_____   _____

Ort, Datum   Nataliia Dmitrieva